Image Quality Analysis

From RAW to Final Image quality

How do you choose a camera?

How do you control camera specifications?

What should you remember from this?

- Know how to read and request raw camera specifications for your applications:
  - Available range of settings: Exposure time, gain, aperture
  - Methods for measuring: Noise/raw capture performance, sharpness and color
  - Useful metrics for summarizing these measurement results
- Know how to do the same for tools & hardware that process RGB images:
  - Evaluation environment: viewing conditions, light levels, presence of motion,…
  - Metrics and methods for measuring: Perceptual noise, sharpness and color
- More generally, have a rough idea of:
  - The landscape of camera performance evaluation methods & standards
  - Their evolution, and why we sometimes need more complex methods

Today’s agenda

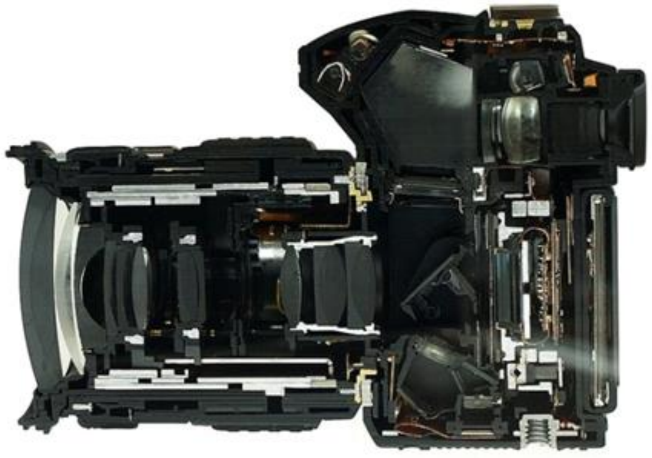
- Comparing RAW sensor performance (applications for hardware preselection):
  - ISO speed rating and sensor gain
  - Exposure time / Timing
  - Sensor noise / SNR curve
  - Sharpness: PSF / MTF
  - Color response: Relative sensitivity, WB scales, color matrix
- Comparing performance in RGB images (applications in photography/human visible images):
  - CIE-L*a*b* color space, perceptual uniformity and color fidelity
  - Viewing conditions
  - Detail preservation: Texture MTF & Acutance
  - Visual noise
- Conclusion

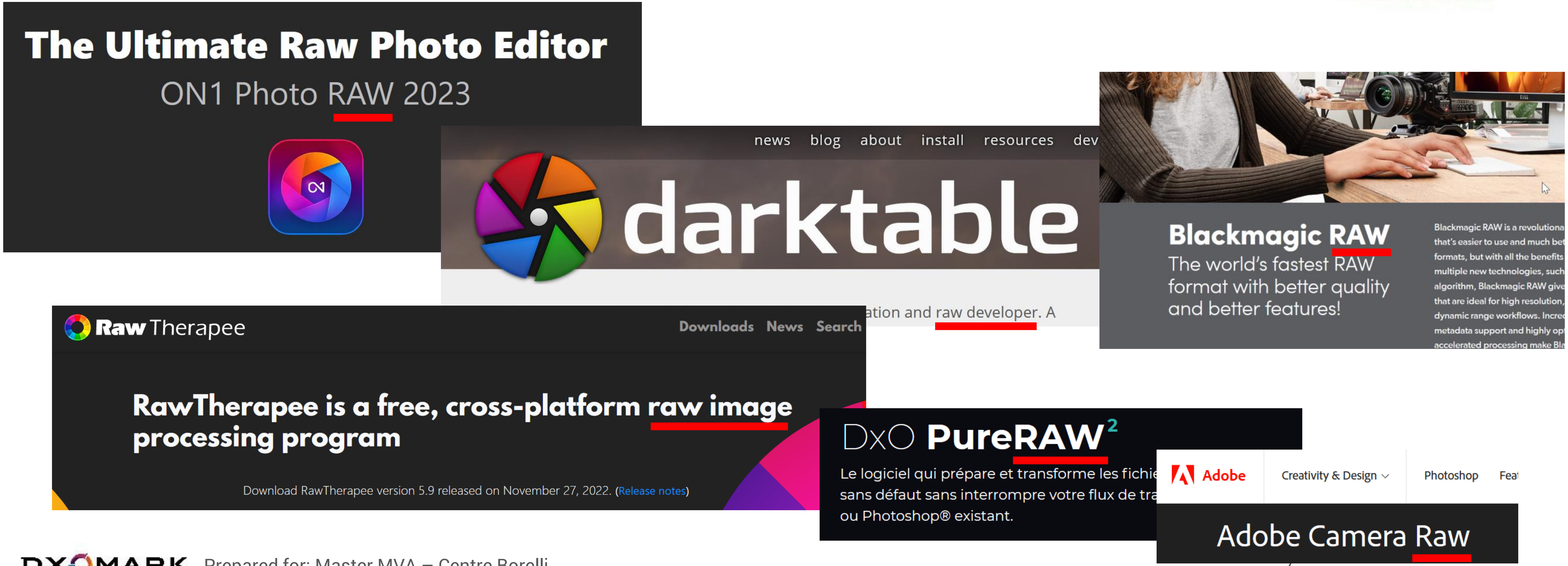
Comparing raw sensor performance

A short recap on linear raw formats

RAW image

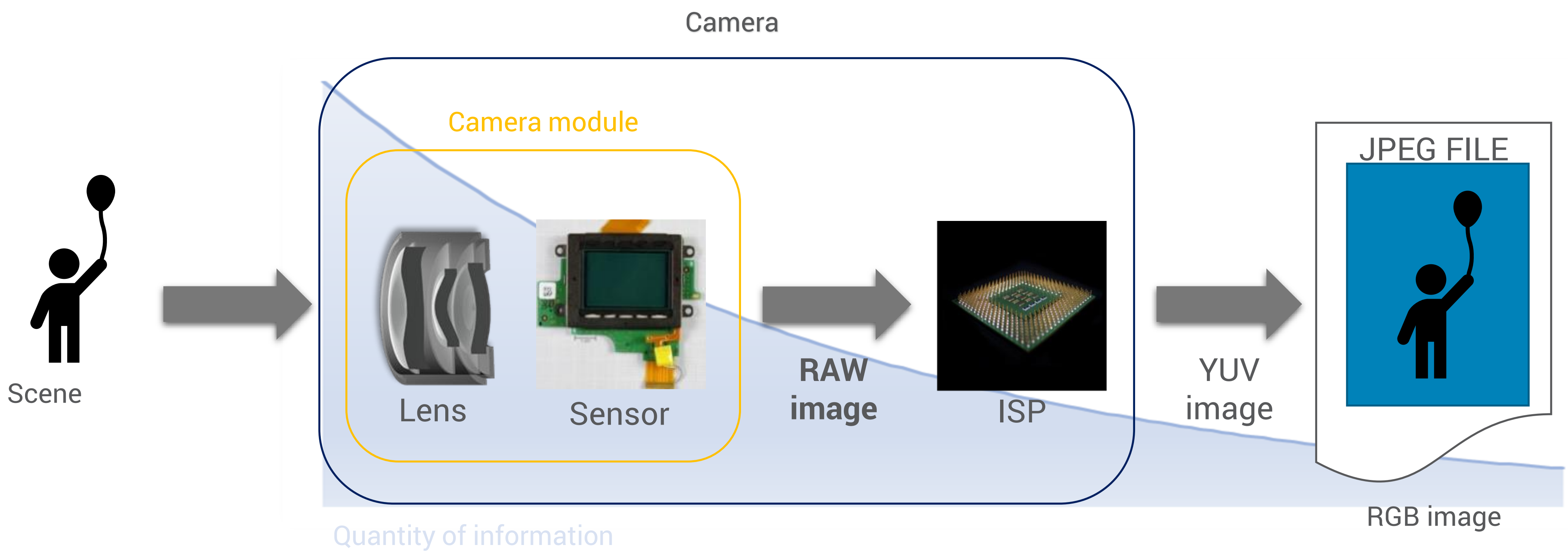
- Why do professional photographer brag about “shooting in RAW”?
- What does it mean?
- And why would it be more “professional”?





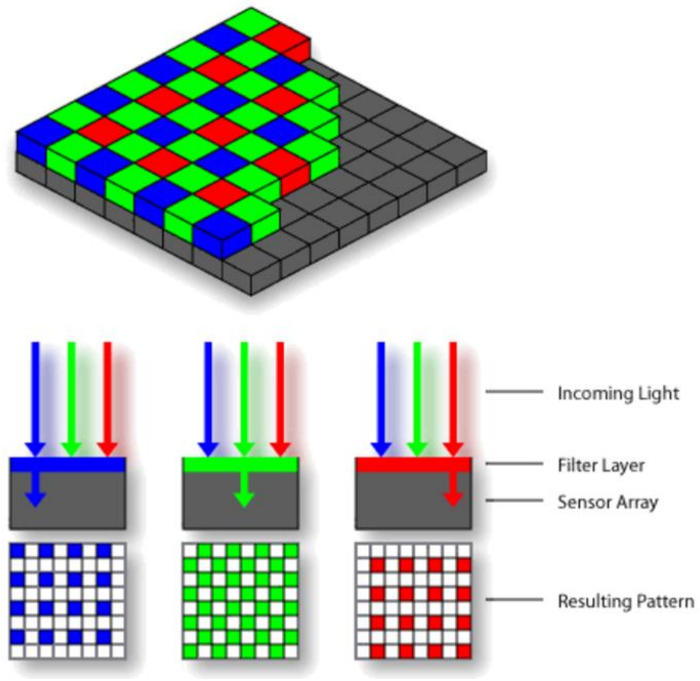
RAW image

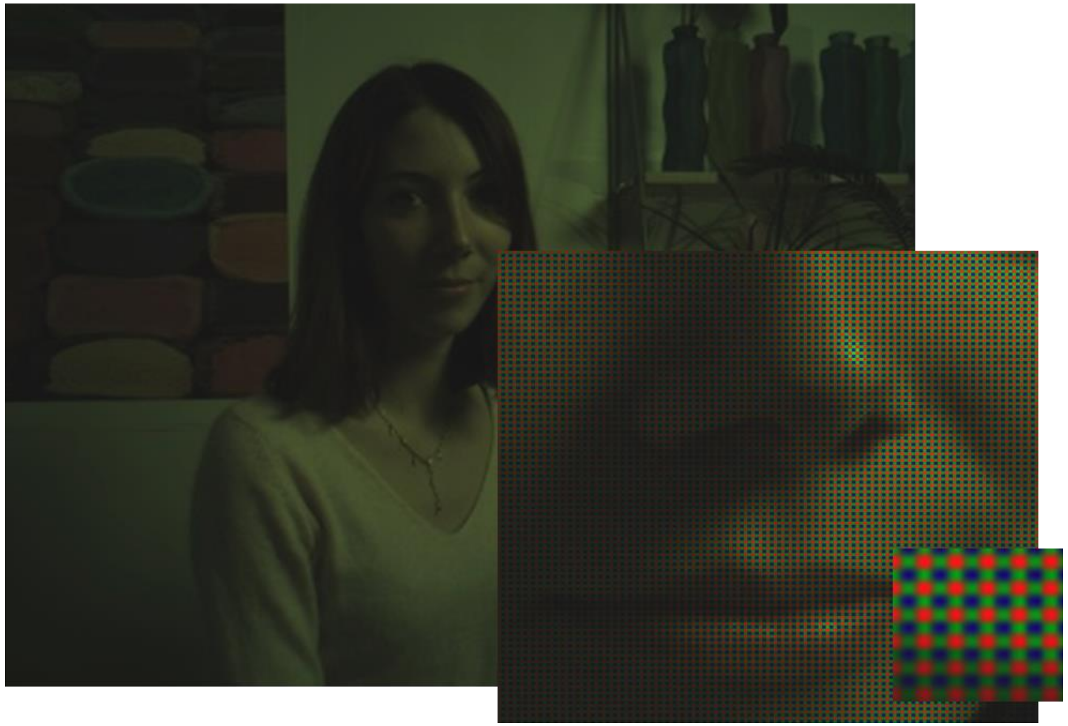
- A RAW image is a snapshot of the state of the sensor with little or no processing applied
- Each processing step will irreversibly reduce the amount of information



RAW image

- It is barely an image in the sense that it is not intended to be viewed by the user
- RAW formats often reflect the structure of the Bayer filter of the sensor (one channel per pixel)



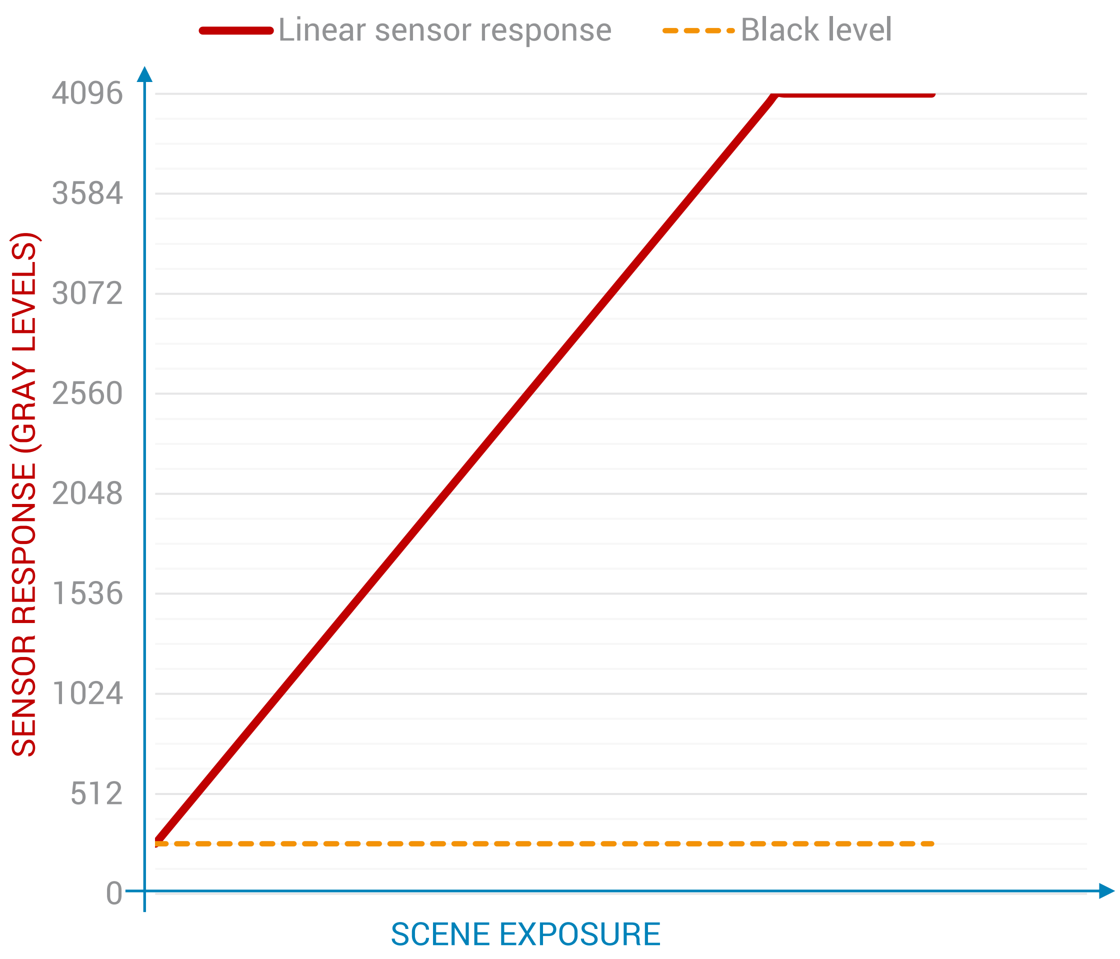


Linear response

- Ideal sensor response (gray level) to light intensity is linear (with an offset)

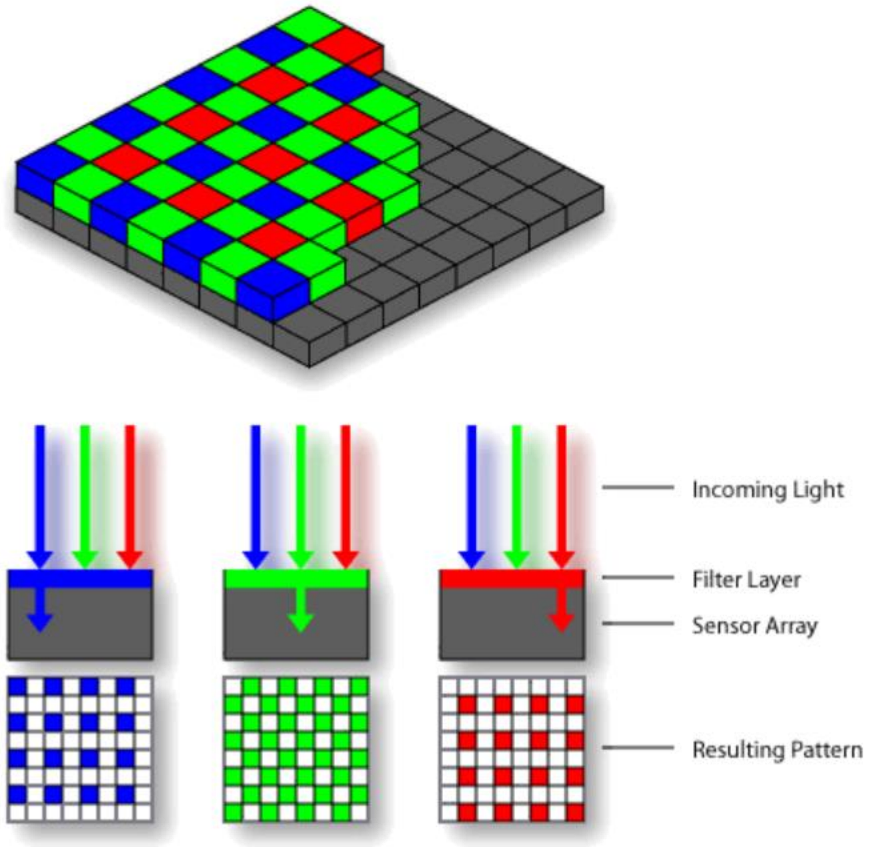
$$GL = \boldsymbol{K}_{\boldsymbol{sensor}} \times N_{photons} + GL_0$$
$$GL = \boldsymbol{K'}_{\boldsymbol{sensor}} \times L_{scene} + GL_0$$

- Useful model:
  - Simple relationship between physical scene exposure and pixel values
  - All early ISP blocks are in linear domain
  - Sensor RAW files are often assumed linear



Bayer filter

- The Bayer Filter filtered the light by wavelength range with a mosaic color filters placed over the pixel sensor to capture color information
- 25% of blue, 25% of red and 50% of green because human's eyes are more sensitive to green light

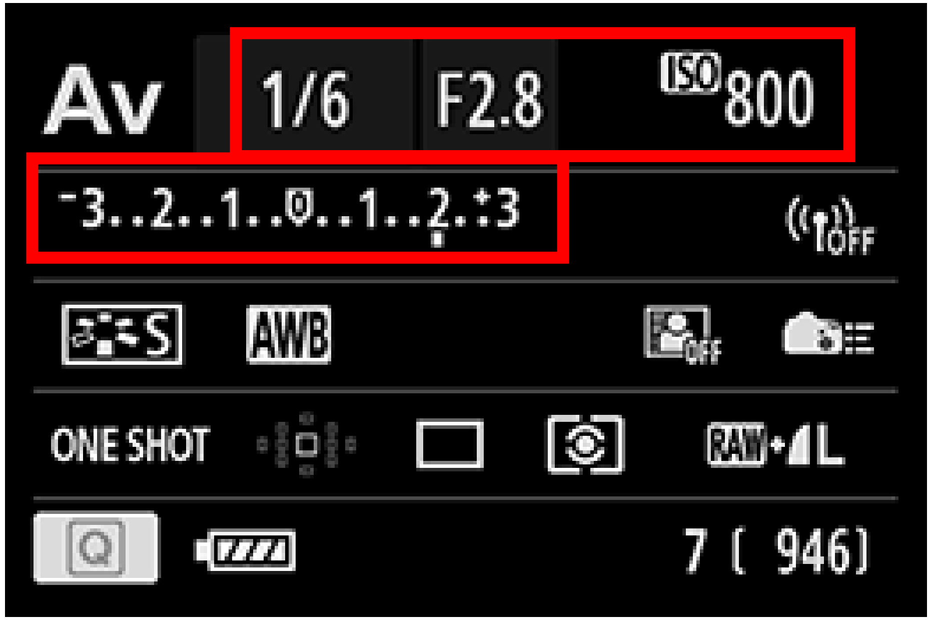


Comparing raw camera performance

How to measure comparable conditions?

Camera settings in consumer products

Camera exposure settings

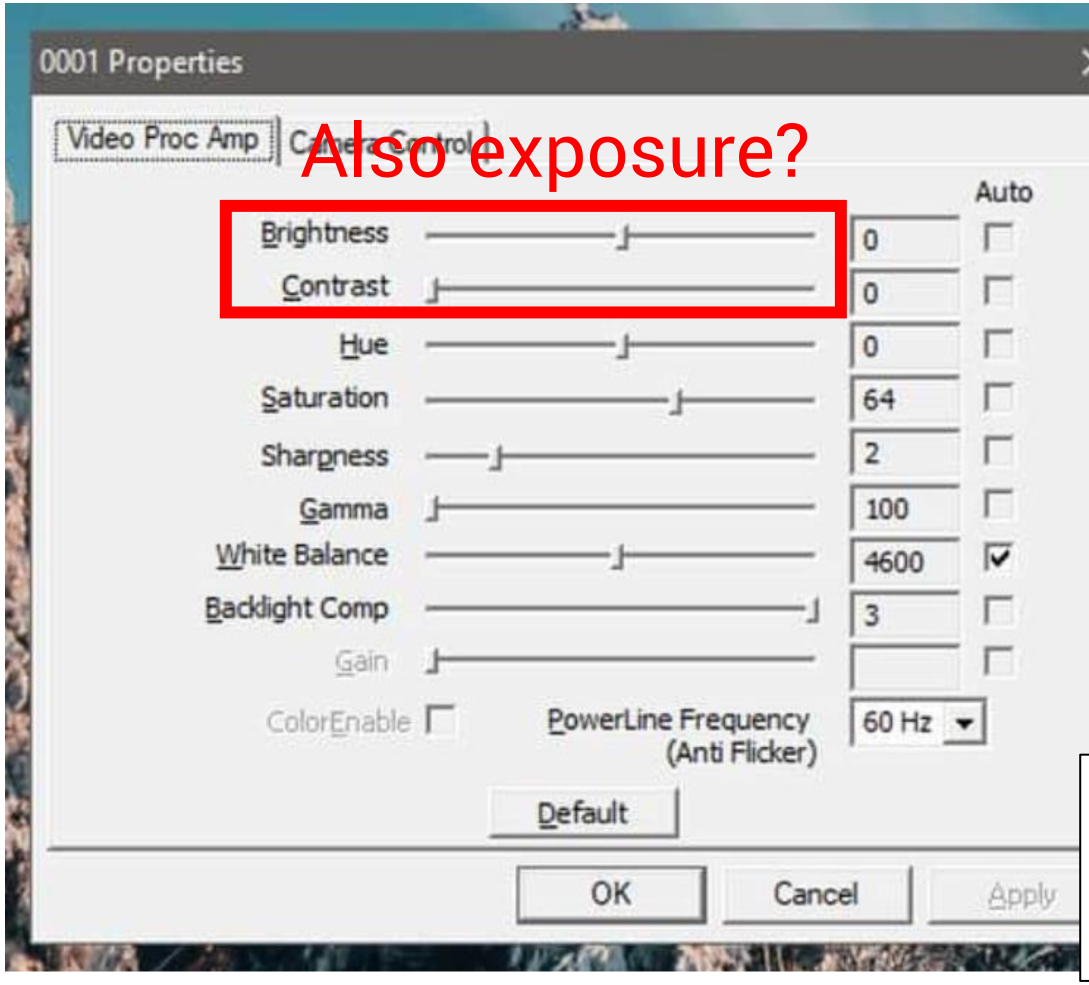


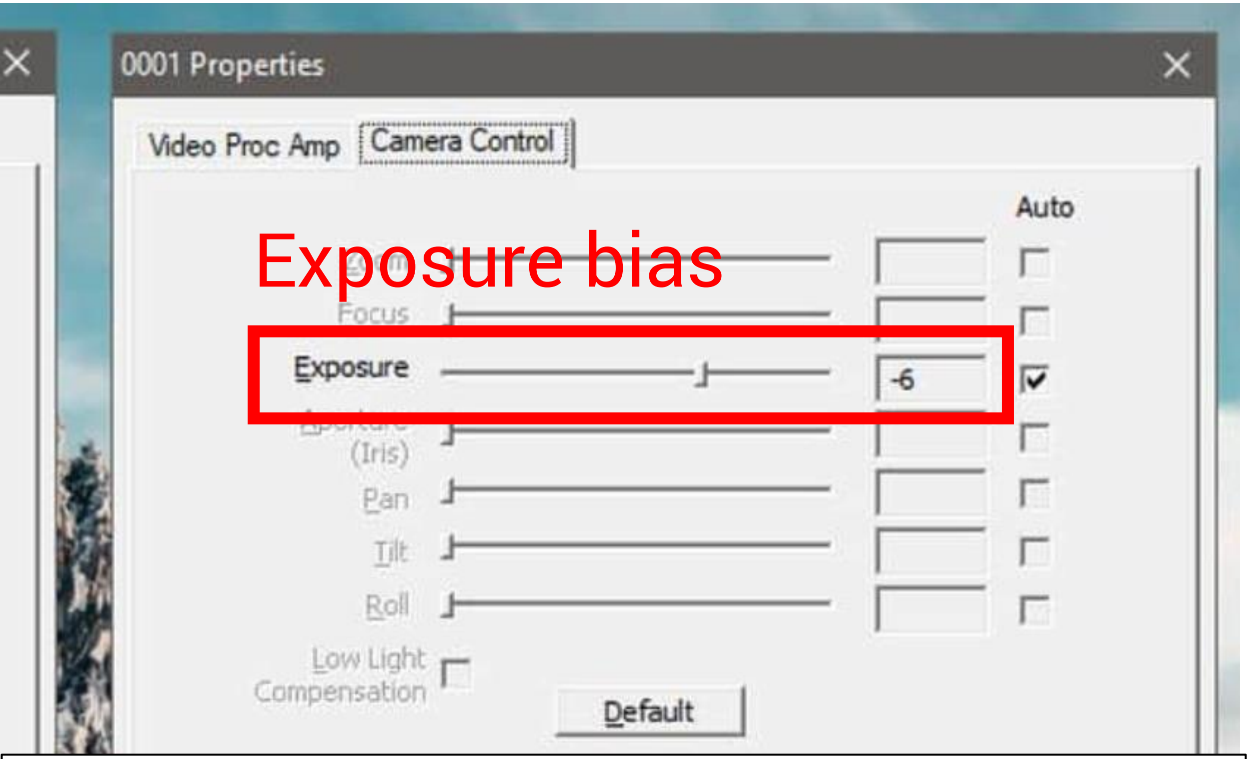
Top, from left to right: exposure time, aperture, ISO.
Below: exposure bias / meter.

Smartphone exposure bias control

Bias "changes" the target level of AE (auto-exposure) algorithms

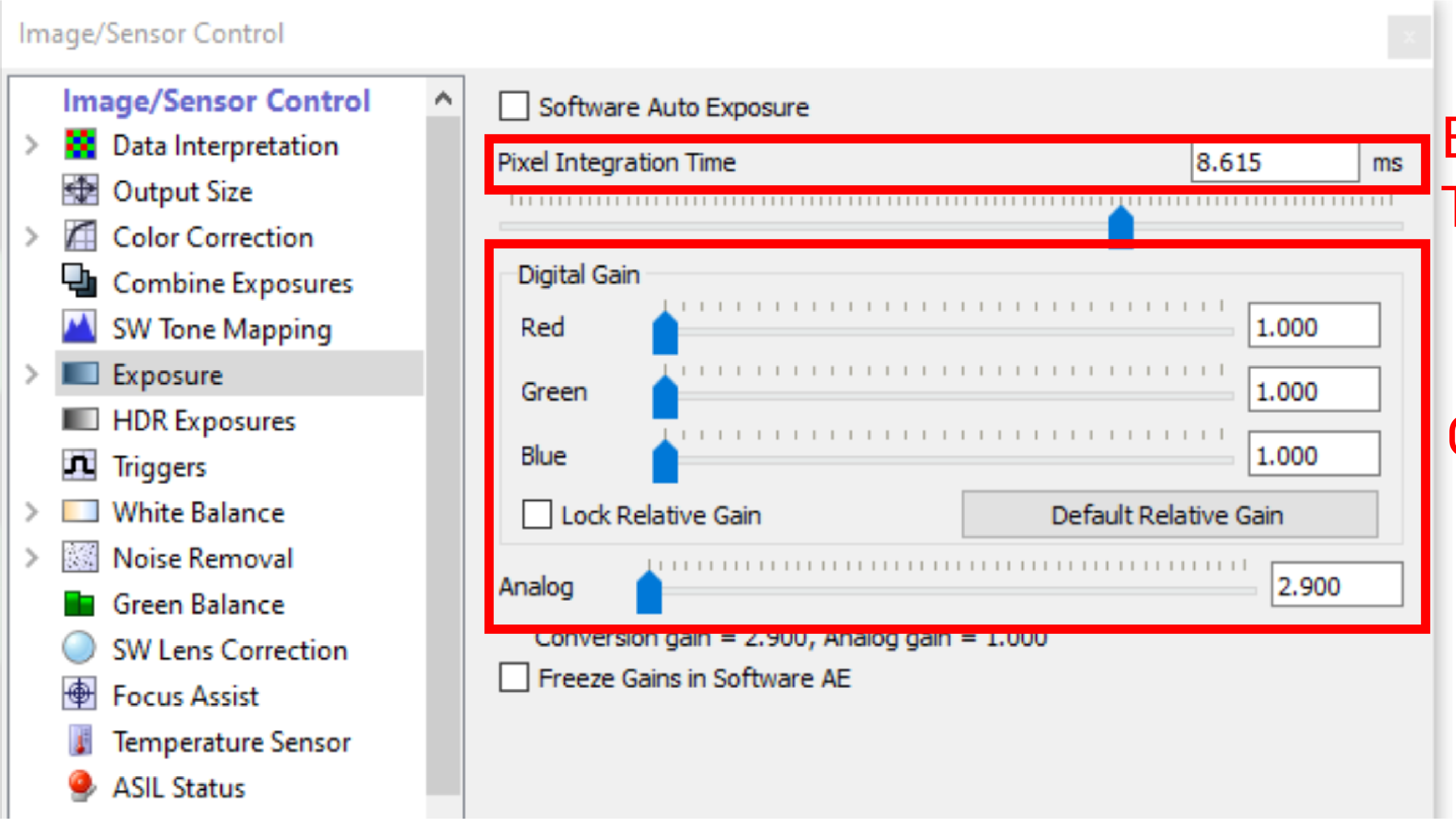
Camera settings in specialized equipment (UVC* cameras)





** UVC (USB Video device Class) is a standard for simple cameras exposed as USB devices, initially designed for webcams but also offered by a lot of industrial devices for convenience. They are generally directly supported by frameworks like OpenCV/Matlab.*

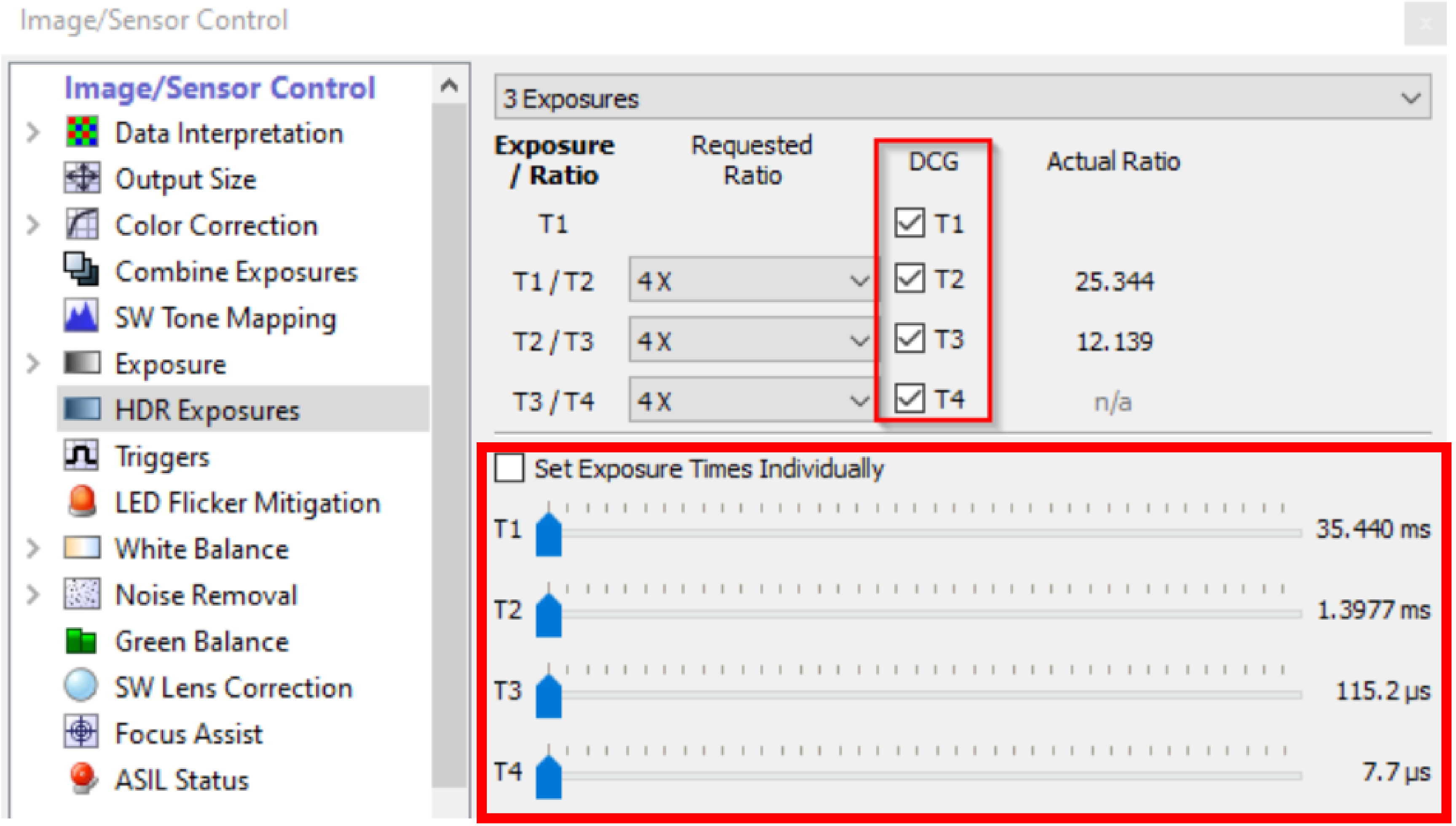
Camera settings in specialized equipment (dev boards)



Exposure Time

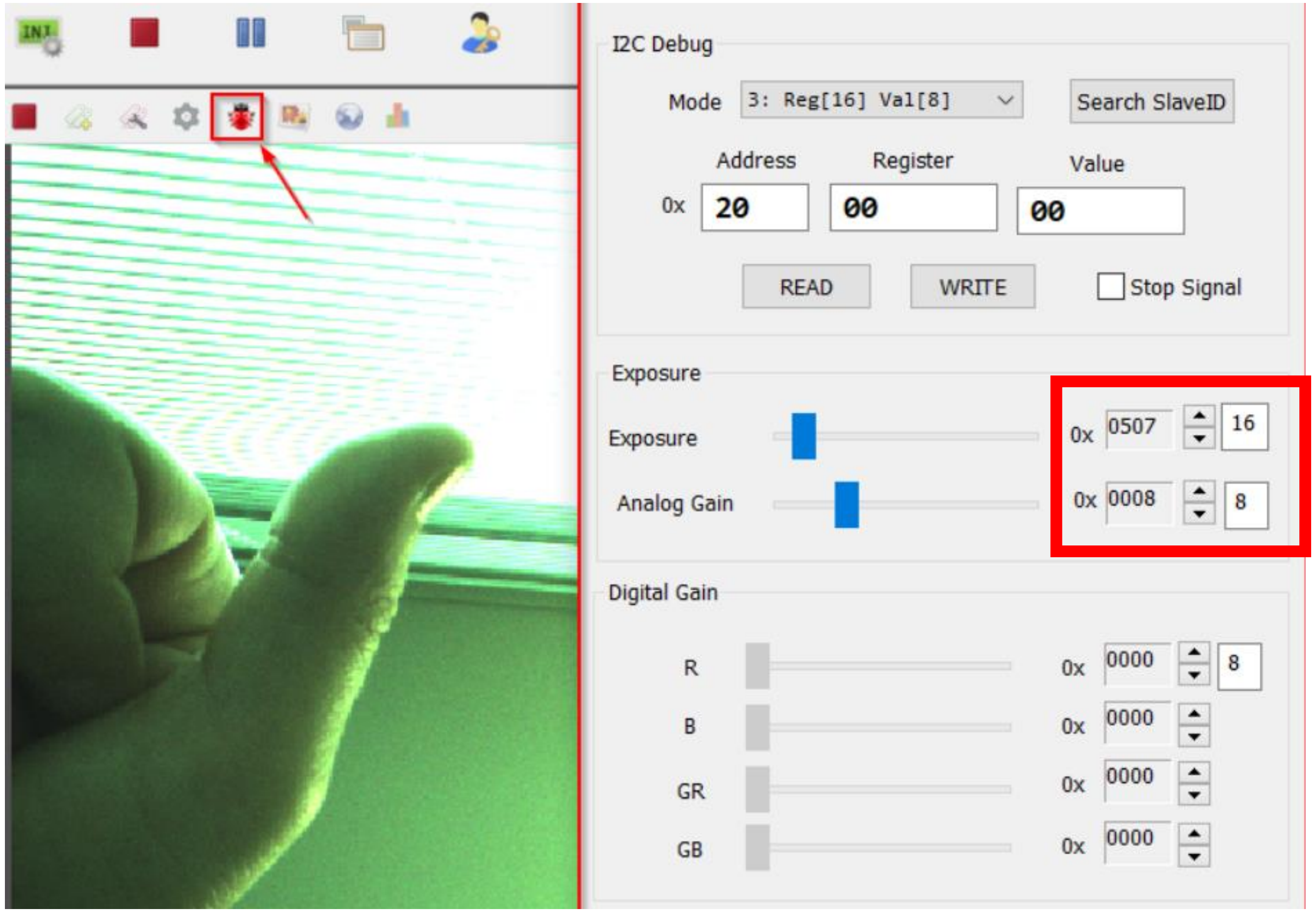
Gains

Camera settings in specialized equipment (dev boards)



Exposure Times

Camera settings in specialized equipment (dev boards)



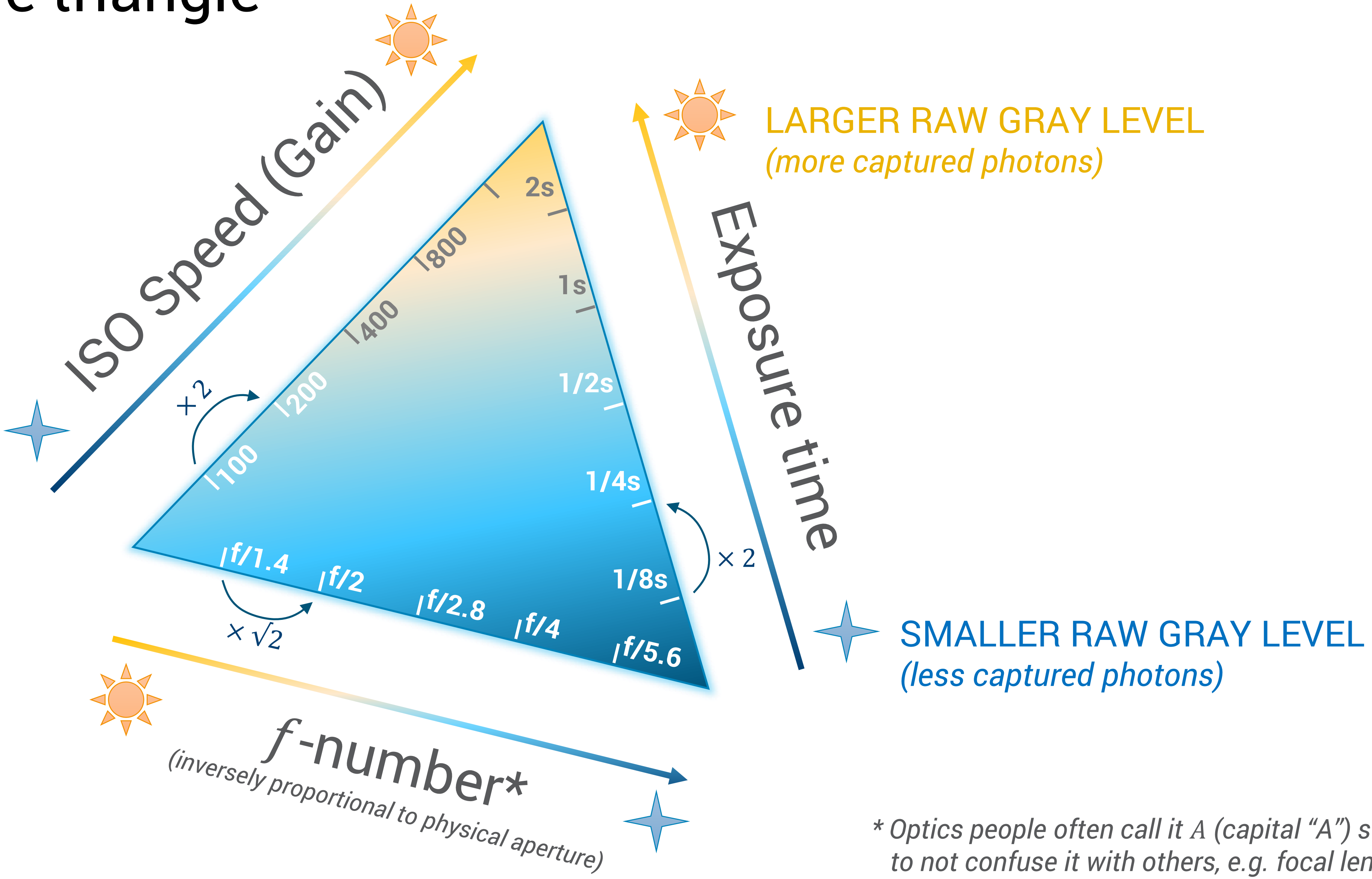
Exposure Register
Gain Register

Camera exposure settings

How to compare devices when all these settings affect gray level values?

- Exposure time(s)
- Exposure bias
- Aperture (sometimes fixed)
- ISO
- System gain(s)
- Low-level registers for exposure, gain,…
- ISP parameters: Brightness, Contrast, LTM,…

The exposure triangle



* *Optics people often call it A (capital "A") so as to not confuse it with others, e.g. focal length.*

Importance of comparing comparable things

- Sensors are best operated in **Photon Shot Noise** regime.

- An ideal sensor that counts photons perfectly with expected value N has:

$$\text{Ideal SNR} = \frac{N}{\sqrt{N}} = \sqrt{N} = C(L)\sqrt{\frac{t}{A^2}}$$

t ← Exposure time (s)

A ← f-number

- A real sensor is able to capture $N/\boldsymbol{k}$ photons in the same conditions:

$$\text{Real SNR} = \frac{N/k}{\sqrt{N/k}} = \sqrt{N/k} = \frac{C(L)}{\sqrt{\boldsymbol{k}}}\sqrt{\frac{t}{A^2}}$$

$\sqrt{\frac{t}{A^2}}$ ← Fixed to find $\boldsymbol{k}$

$\sqrt{\boldsymbol{k}}$ ← "Ability to capture photons"

Importance of comparing comparable things

- In other words: a real sensor is able to capture less photons and has to apply a gain $\boldsymbol{k}$ (e.g. ISO) to provide the same output gray level values:

$$\text{Real SNR} = \frac{C(L)}{\sqrt{\boldsymbol{k}}}\sqrt{\frac{t}{A^2}}$$

- Comparing two sensors in PSN regime needs the same conditions:
  - t: Exposure time (more time captures more photons)
  - A: Lens aperture (large aperture captures more photons)
  - $\boldsymbol{k}$: Linear gain required to achieve the same gray levels with t and A.

Comparing raw camera performance

How to measure comparable conditions? – Measuring ISO speed rating & exposure time

ISO speed rating

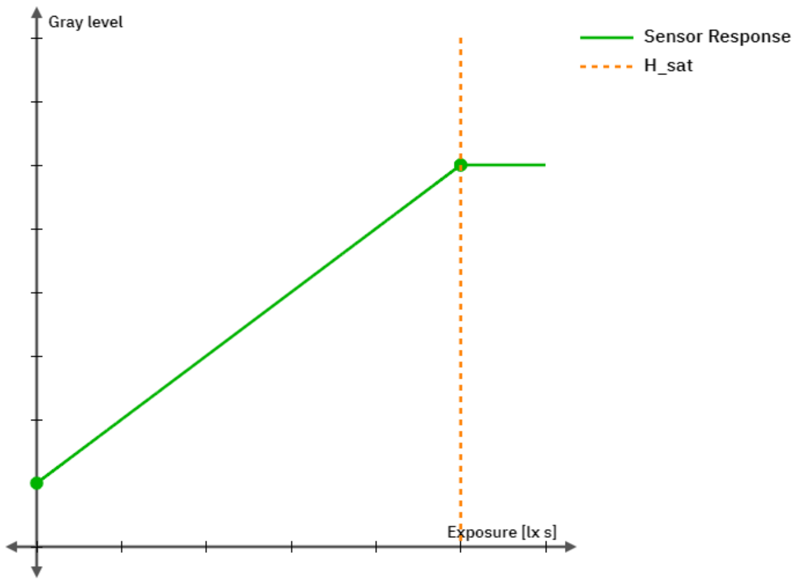
- Knowing the aperture and the exposure time, this allows to compute the sensor response, which gives us the minimal exposure (in $lx.s$) required to saturate the sensor as:

$$H_{sat} = 0.65\, L_{sat}\, t/A^2$$

- The ISO Sensitivity (unitless) is then defined as:

$$S_{sat} = 78\, lx.s/H_{sat}$$

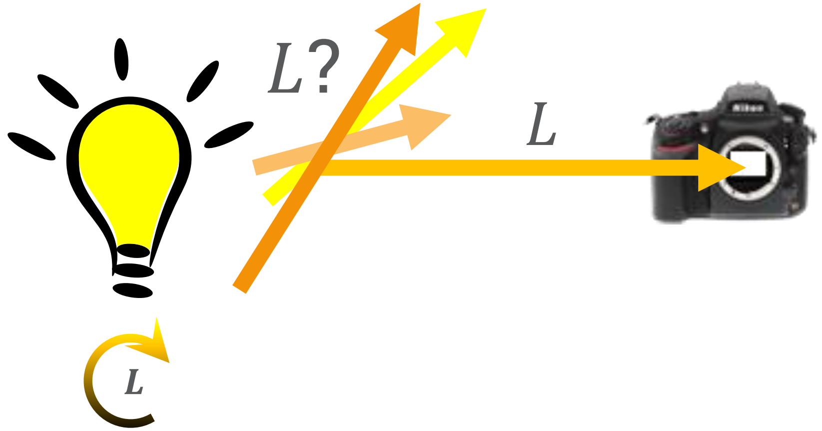
- This is the standard way to express the actual sensor gain in visible photography (usually rounded to 100, 200, 400, etc).

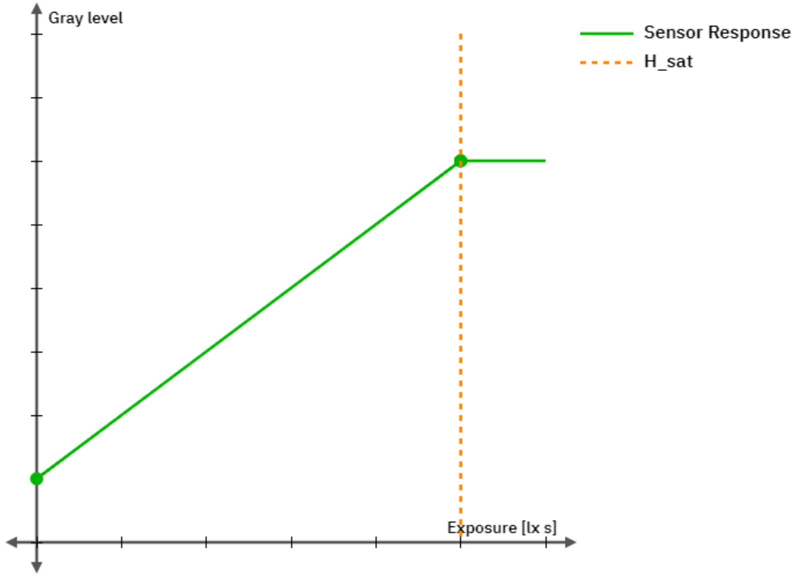


$$\boxed{GL = f_{ISO}(L) = \frac{0.65}{78}\frac{t \cdot L}{A^2} \times ISO}$$

ISO speed measurement

- How to measure $\boldsymbol{L_{sat}}$? With a regular light source?

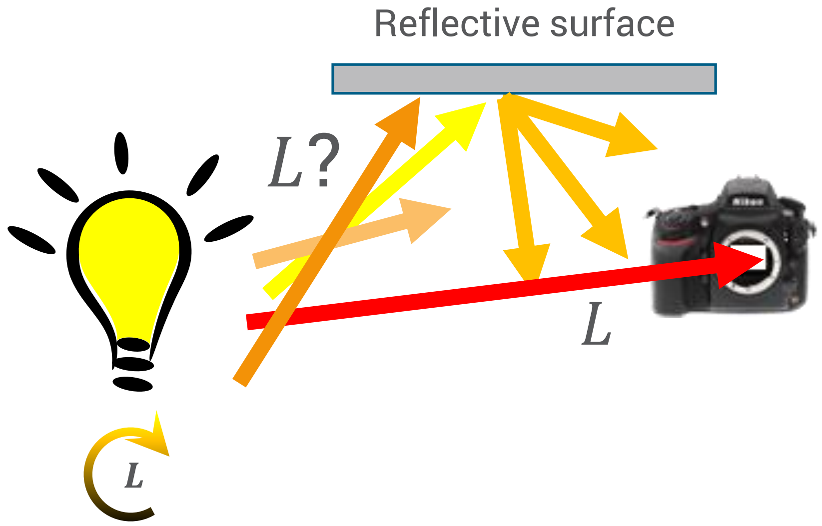


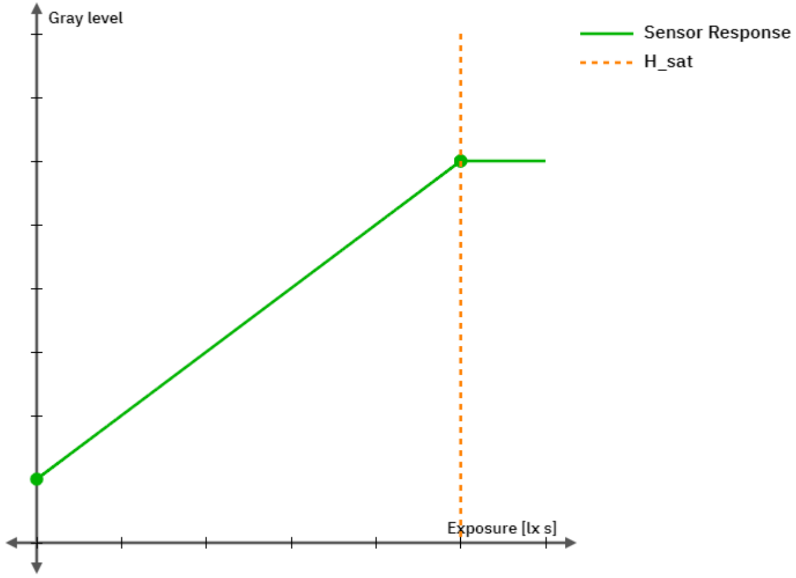


$$GL = f_{\boldsymbol{ISO}}(L) = \frac{0.65}{78} \frac{t \cdot \boldsymbol{L}}{A^2} \times \boldsymbol{ISO}$$

ISO speed measurement

- How to measure $\boldsymbol{L_{sat}}$? On a reflective surface?

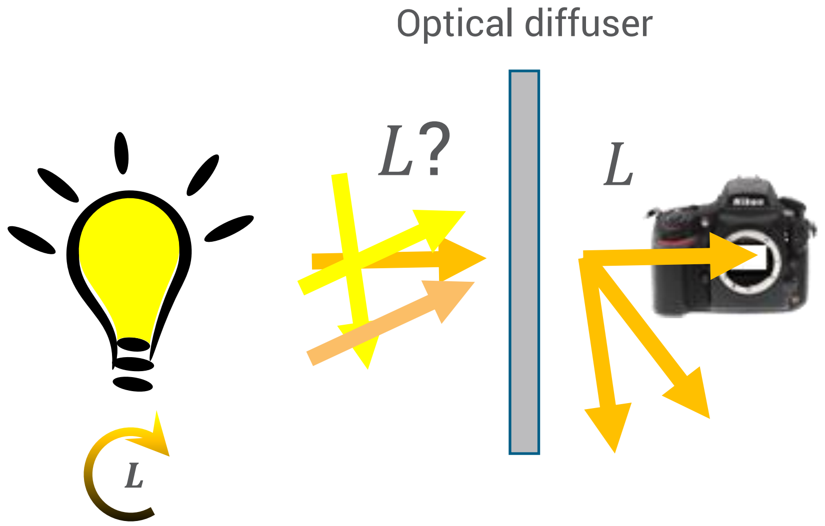


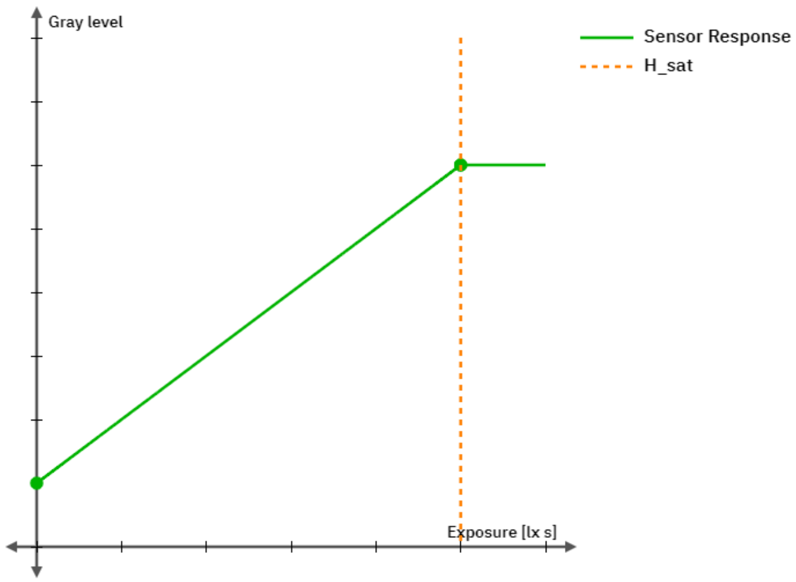


$$GL = f_{ISO}(L) = \frac{0.65}{78}\frac{t \cdot \boldsymbol{L}}{A^2} \times \boldsymbol{ISO}$$

ISO speed measurement

- How to measure $\boldsymbol{L_{sat}}$? With an optical diffuser?

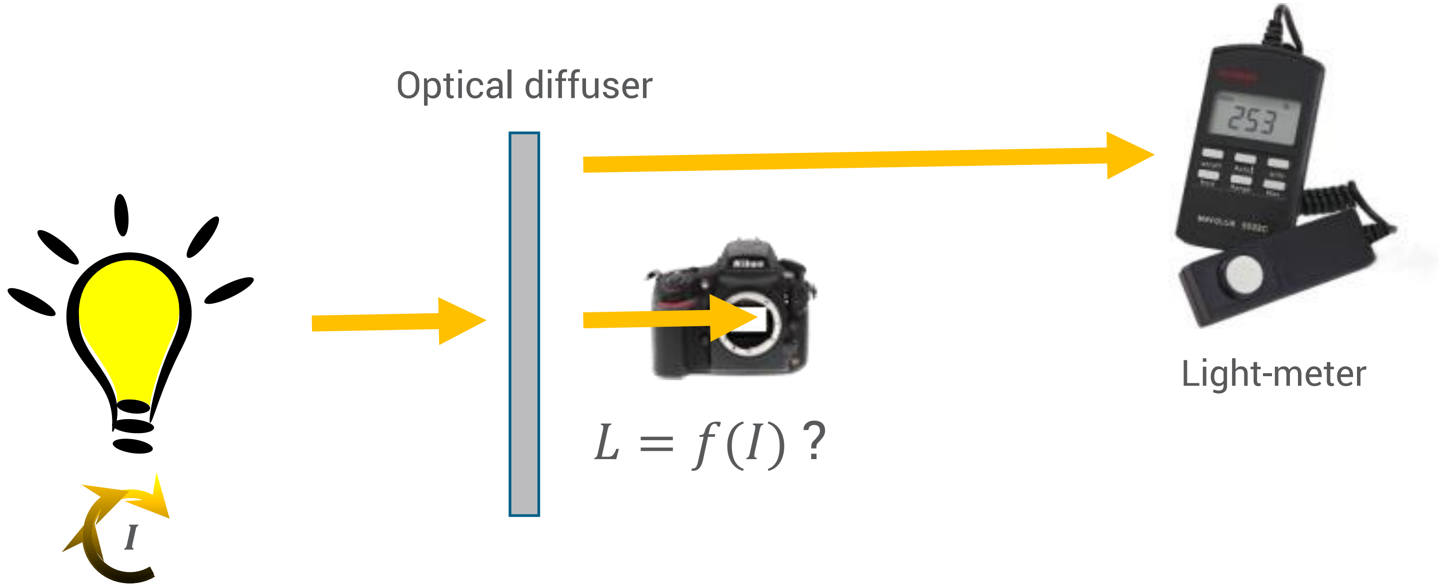


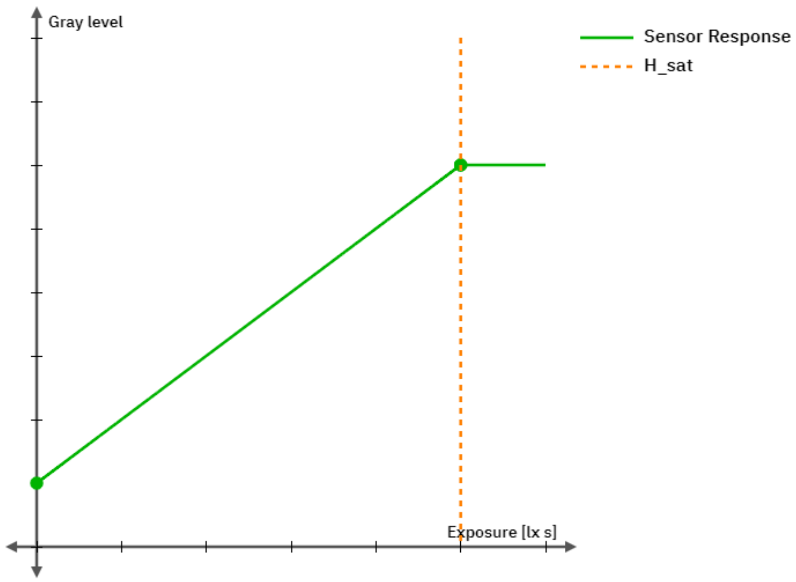


$$GL = f_{ISO}(L) = \frac{0.65}{78}\frac{t \cdot \boldsymbol{L}}{A^2} \times \boldsymbol{ISO}$$

ISO speed measurement

- How to measure $\boldsymbol{L_{sat}}$? In fact, how to measure $\boldsymbol{L}$?



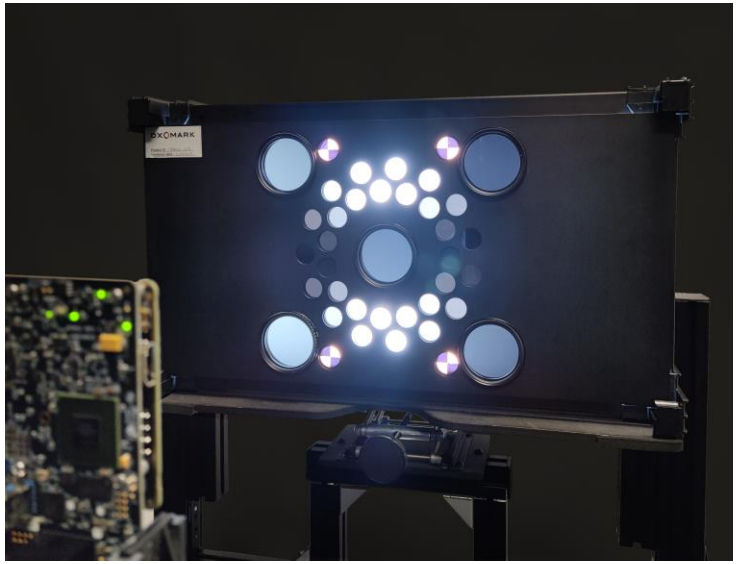


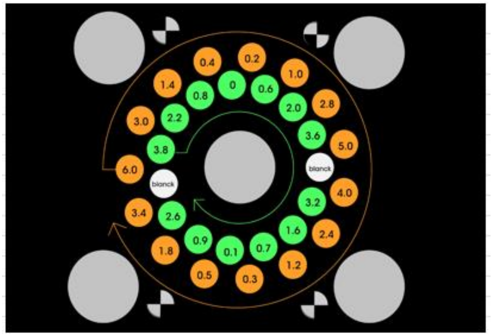
$$GL = f_{\boldsymbol{ISO}}(L) = \frac{0.65}{78}\frac{t \cdot \boldsymbol{L}}{A^2} \times \boldsymbol{ISO}$$

ISO speed measurement – Lab setup

What we actually use:

- LED panel with controllable intensity
- Maximum luminance 25000 cd/m²
- Dynamic range: 120 dB (20 bits)
- 28 neutral density filters
- Avoids imperfections from printing
- Avoids variations in time
- Large target
- No flickering (frequency 250kHz)

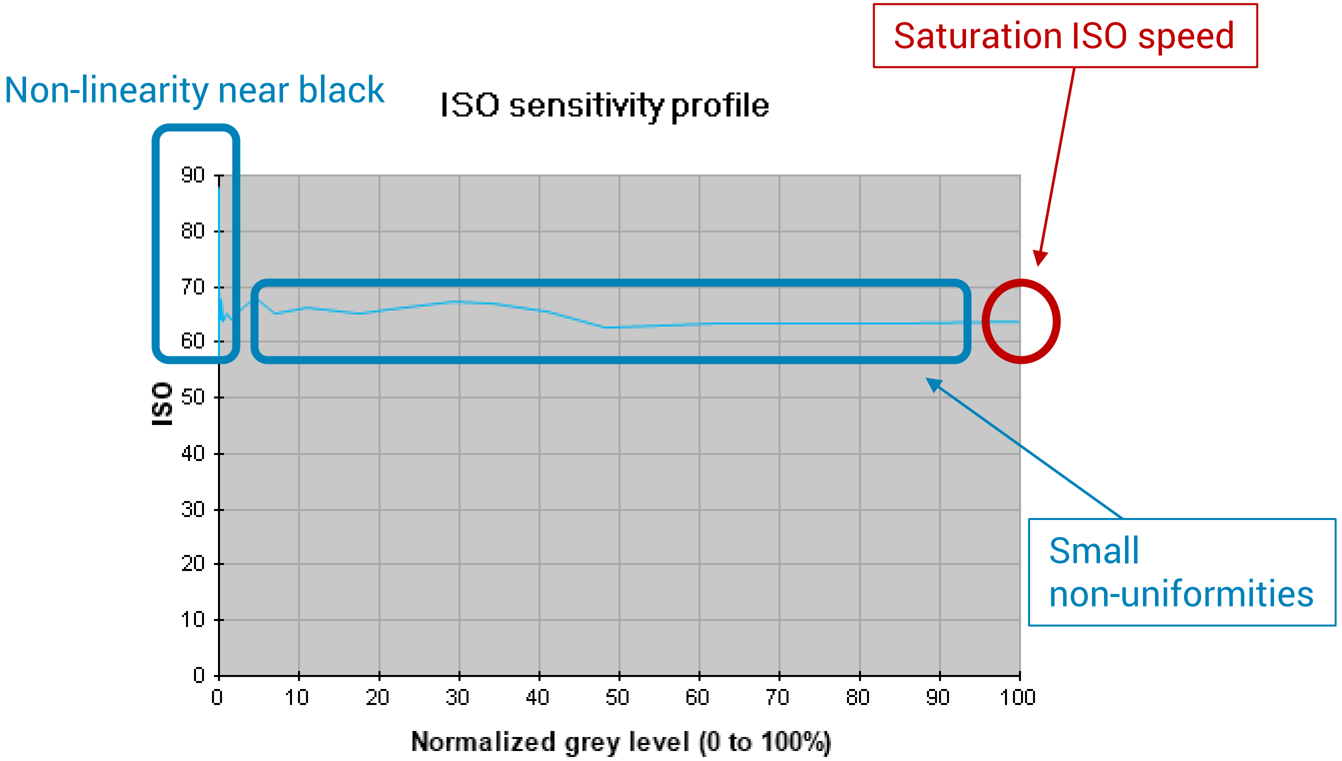




HDR Noise chart

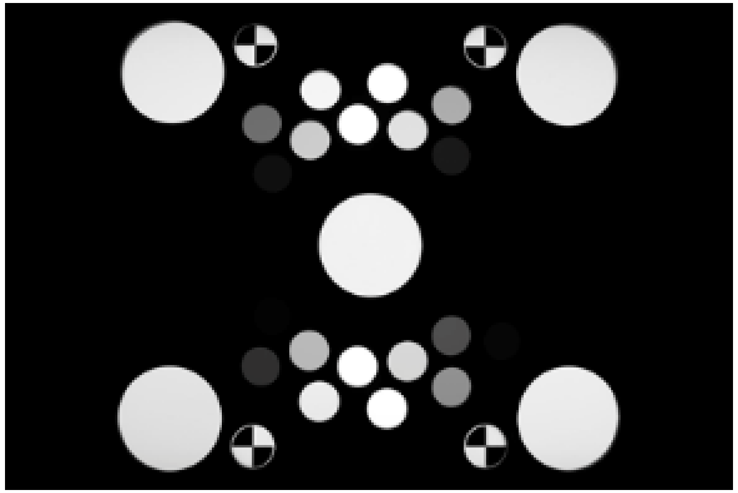
Kino Flo panel

ISO speed measurement – Example output



For multiple patches, trace:

$$ISO = f(GL) = C\frac{A^2}{t \cdot L_{scene}} \times GL$$

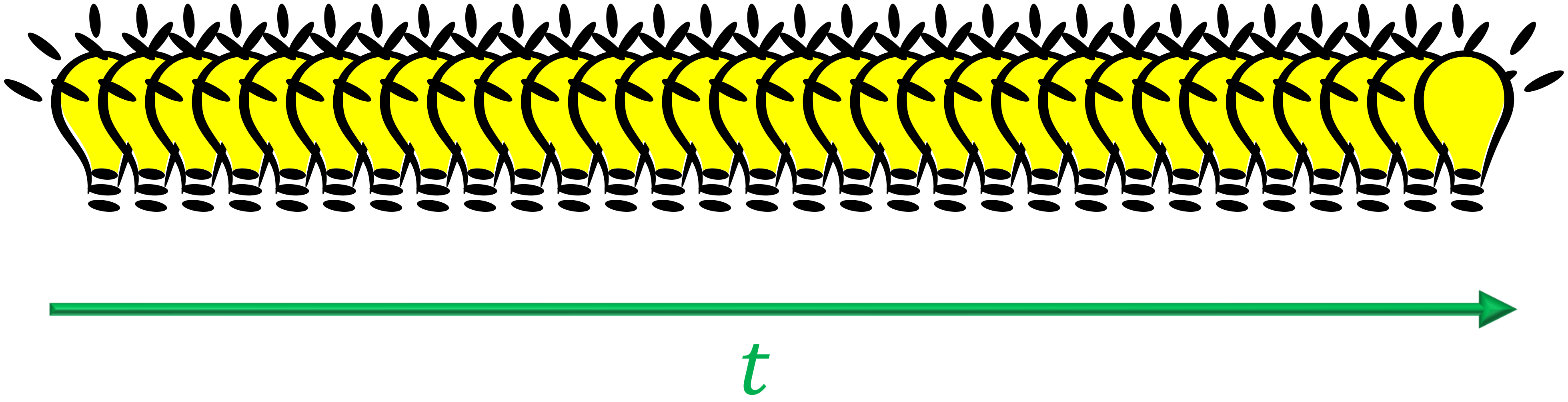


- Circular structure of the target minimizes impact of vignetting
- Non-increasing patch order minimizes impact of flare

Exposure time

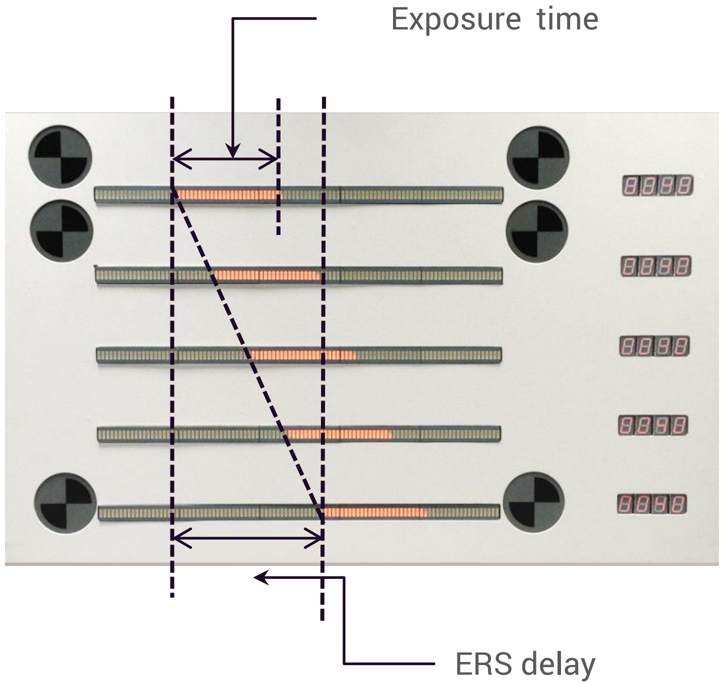
- How to measure t?

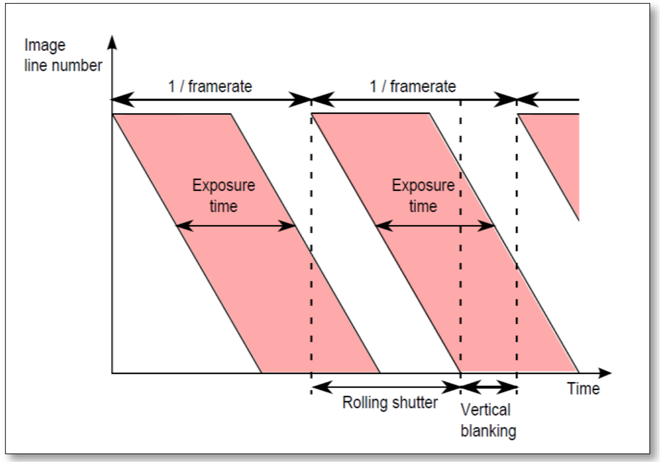
Moving light source = **Motion Blur**!



Exposure time

- One shot of LED Universal Timer processed with Analyzer software provides complete Exposure & ERS (Electronic Rolling Shutter) characterization:





Comparing raw camera performance

Measuring sensor noise

Sensor noise measurements

- A real sensor is able to capture less photons and has to apply a factor k (linked to ISO speed) to provide the same output gray level values:

$$\text{Real SNR} = \frac{C(L)}{\sqrt{\boldsymbol{k}}}\sqrt{\frac{t}{A^2}}$$

- …assuming it is operating in pure photon shot noise regime!

- In real life the sensor always operates in mixed regime.

- Noise performance depends on the entire $SNR = f(N)$ curve (with fixed t, A)

Sensor noise measurements – Noise model

Noise type	Mathematical modeling
Photon shot noise	$\sigma^2_{photon} = k_{PSN} \cdot GL$ Poisson Law – Variance is equal to the expected value (signal)
Dark noise and read noise	$\sigma^2_{dark} = k_{dark}$ Variance is independent of the signal
Photo-Response Non-Uniformity (PRNU)	$\sigma^2_{prnu} = k_{prnu} \cdot GL^2$ Variance is proportional to the signal squared

≈ Gaussian (for large N)
Same for all sensors!

Not Gaussian,
sensor dependent

Not Gaussian,
sensor dependent

$$\sigma^2(GL) = k_{pnru} GL^2 + k_{PSN} GL + k_{dark} \rightarrow SNR(GL) = \frac{GL}{\sqrt{k_{pnru} GL^2 + k_{PSN} GL + k_{dark}}}$$

Sensor noise measurements – SNR curve

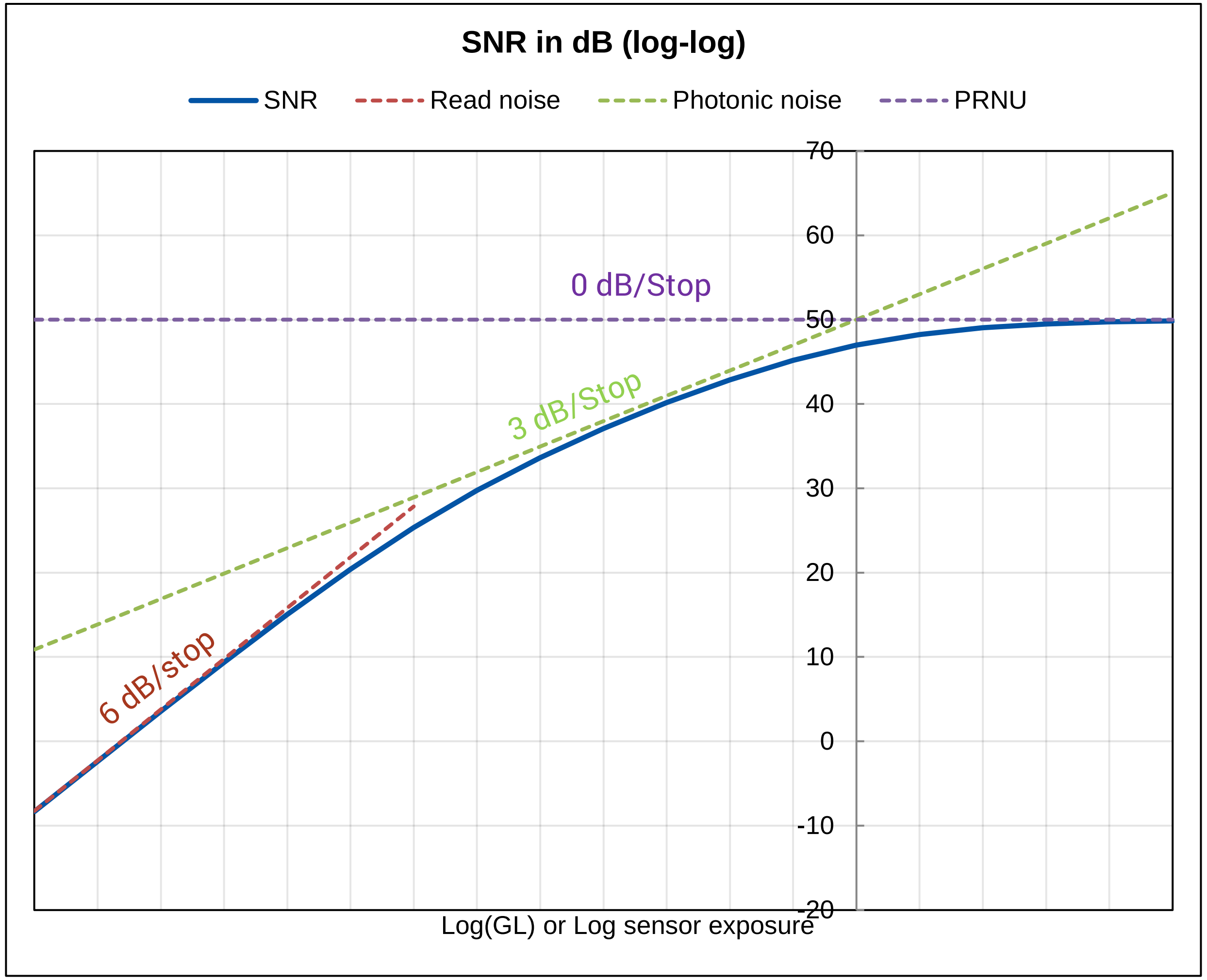
$$SNR(GL) = \frac{GL}{\sqrt{k_{pnru} GL^2 + k\ GL + k_{dark}}}$$

3 modes for RAW SNR:

Shadows: Read noise

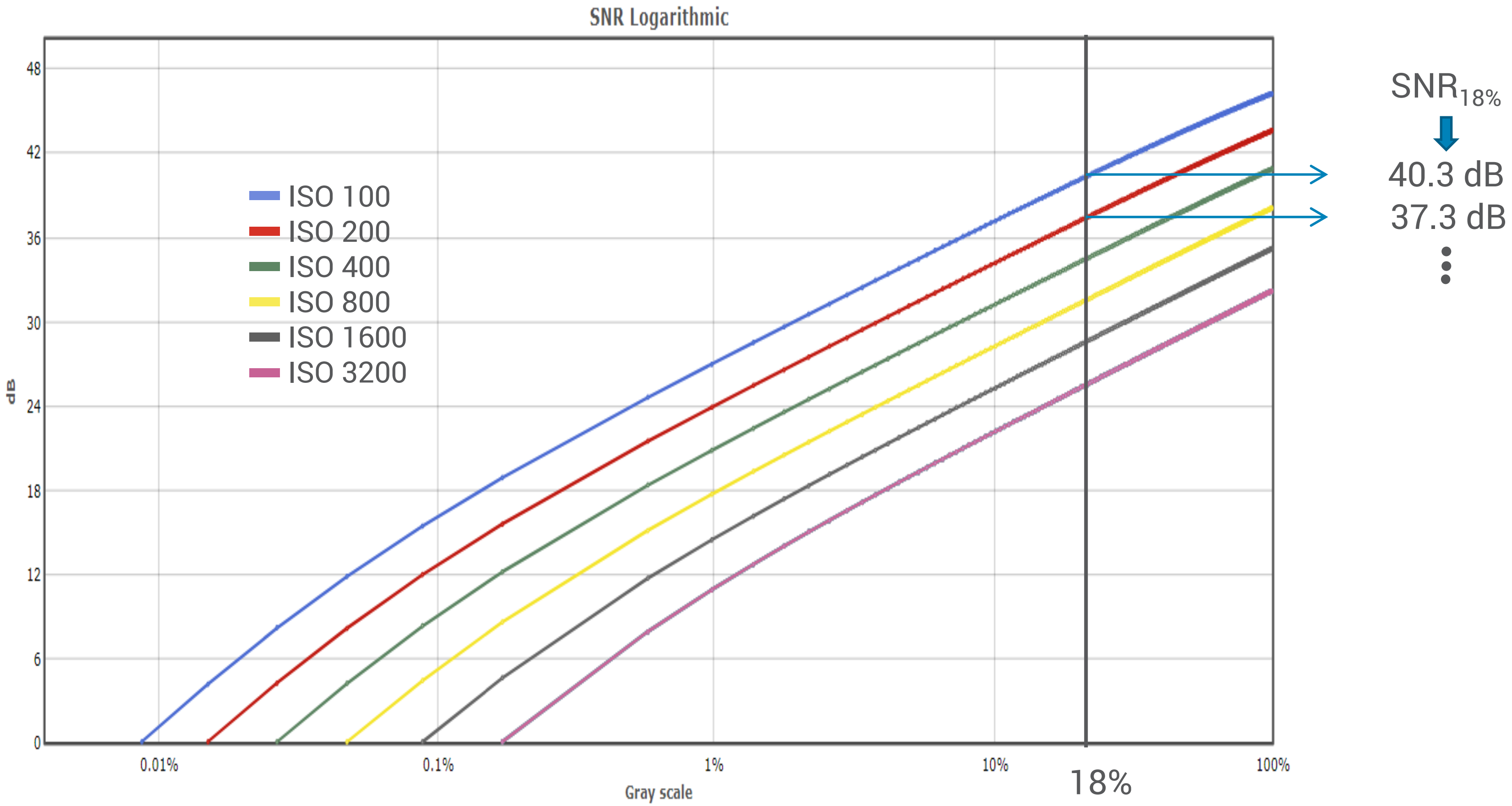
Mid-tones: Photon shot noise

Highlights: PRNU

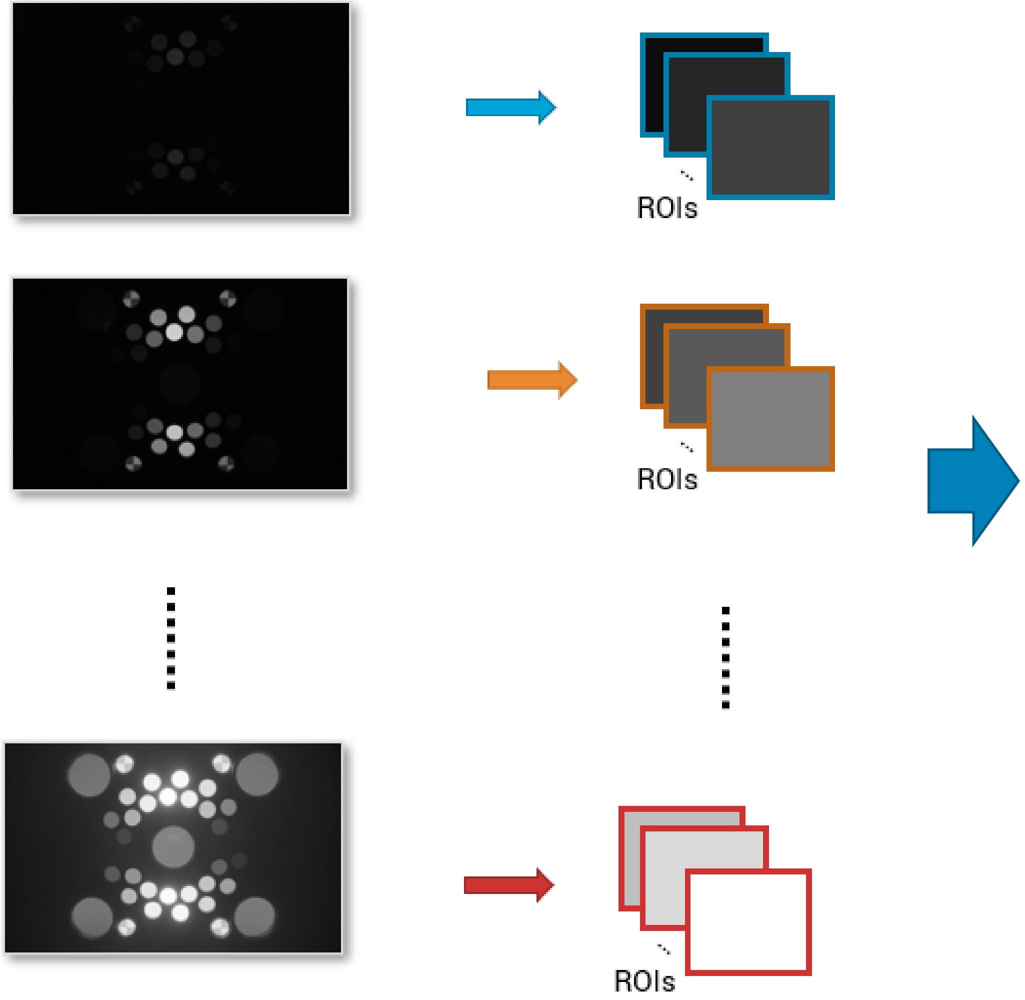


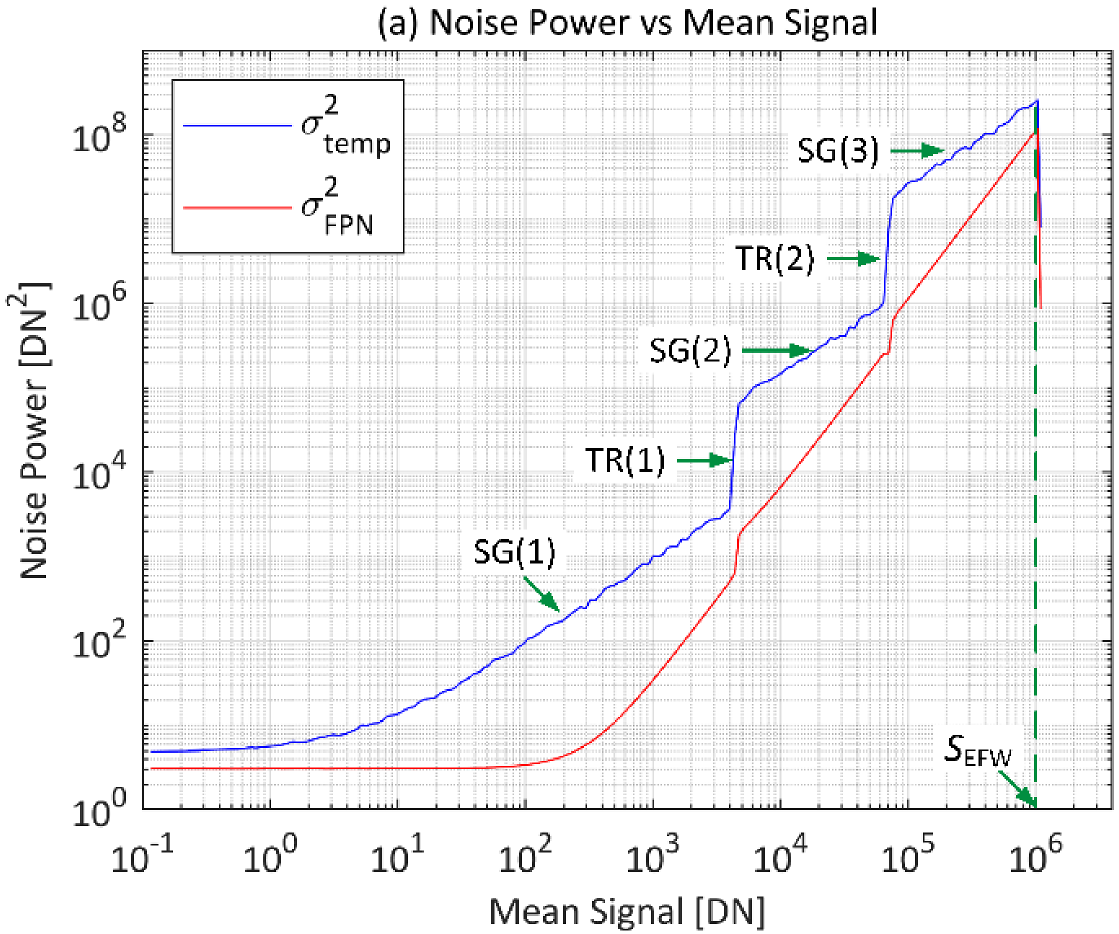
Sensor noise measurements – SNR curve related to ISO/gain

- A high ISO speed implies a shorter exposure time for the same exposure level of the sensor
- Consequently, fewer photons are captured by the sensor and the SNR is degraded



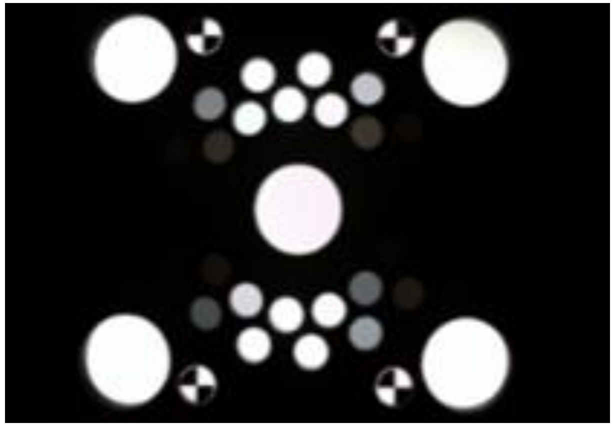
Sensor noise measurements – Multi-Exposure aggregation





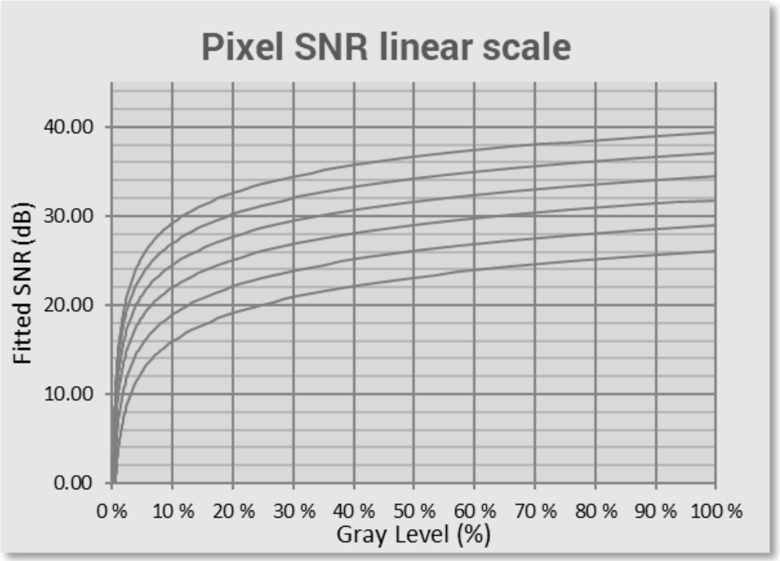
Log-Log Noise Power curve in multi-exposure HDR sensor
Figure from IEEE-P2020 Noise Draft (not published yet)

Sensor noise evaluation in PSN regime – SNR 18%



Some simple performance metrics are based on the SNR at a specific gray level value in photon shot noise regime:

- Shoot the target chart on a backlit panel



- Signal level is measured with the average gray level of each patch
- Noise level is measured with the standard deviation of each patch

- The Signal-to-Noise ratio is simply defined as the ratio between the two, in decibel (dB).

- The SNR can be fitted for signal levels not present in the measured image.

- The SNR 18% is simply the SNR at 18% of the sensor dynamic.

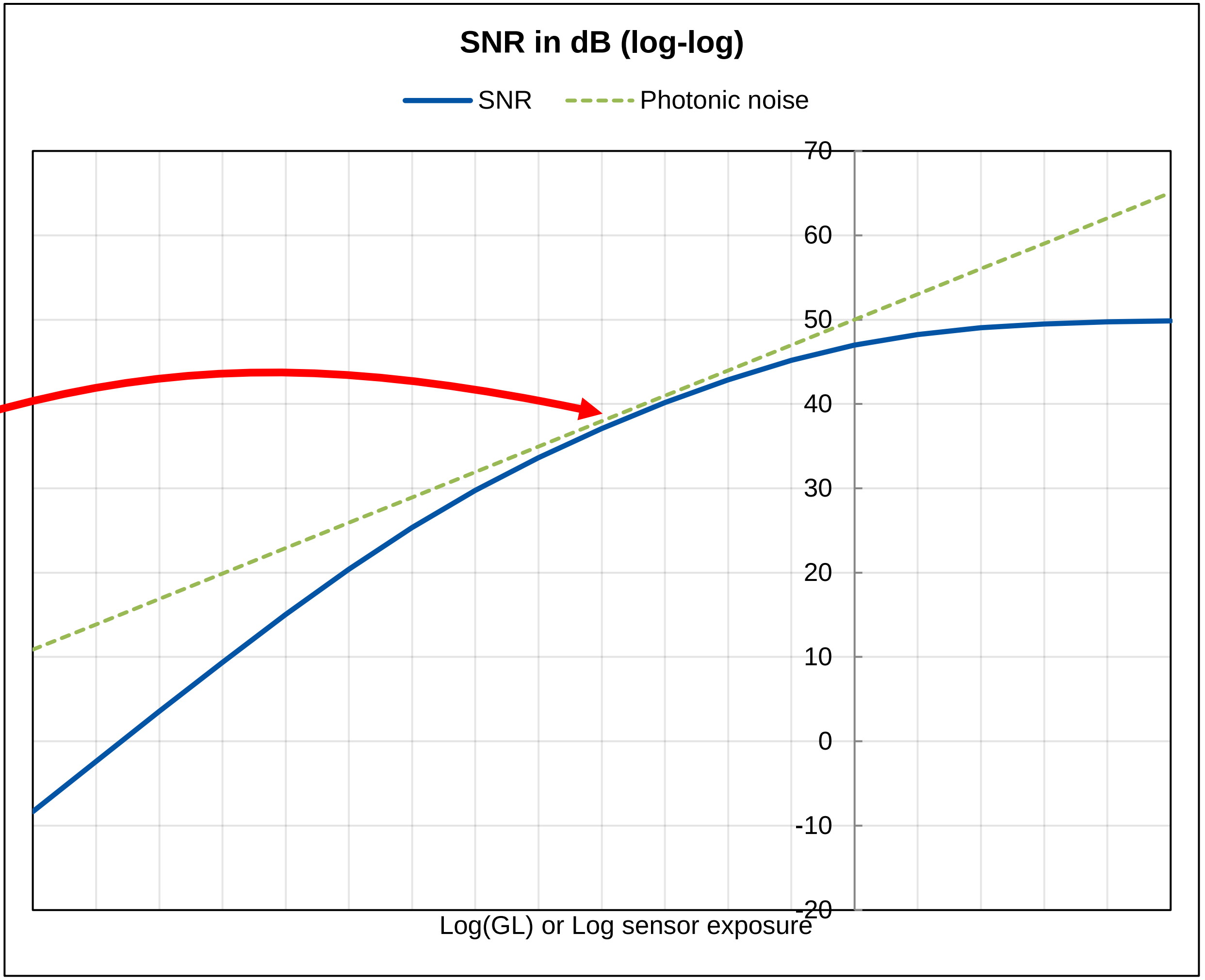
Sensor noise evaluation in PSN regime – System Gain

Shot Noise (SN): dominant where the noise power is linear with the mean signal (slope = 1 in log-log scale)

The System Gain is defined in GL/e^{-} as:

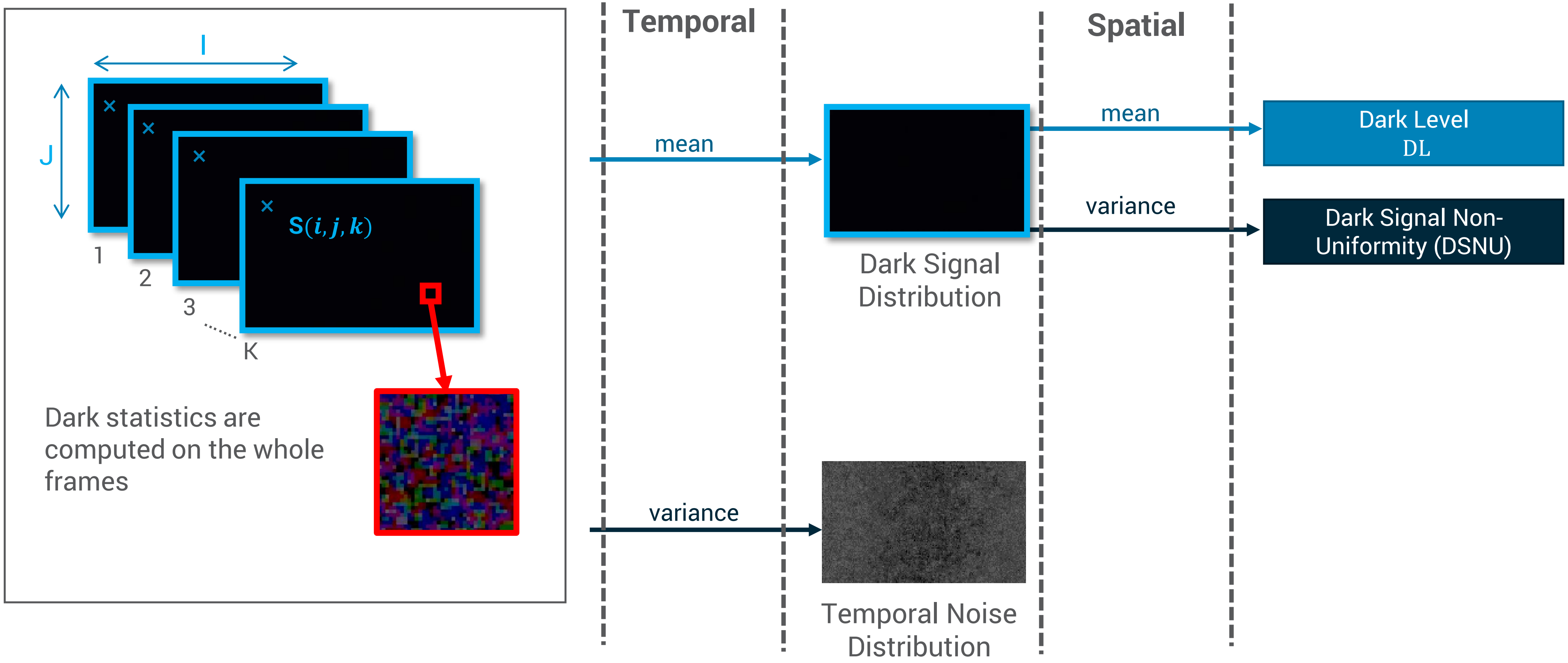
$$\sigma_{\mathrm{SN}}^{2} = \mathrm{SG} \cdot \mu_{\mathrm{SN}}$$

(System gain is included in our constant ***k*** from earlier!)



Formula from IEEE-P2020 Noise Draft (not published yet)

Sensor noise evaluation in read noise regime – Dark statistics



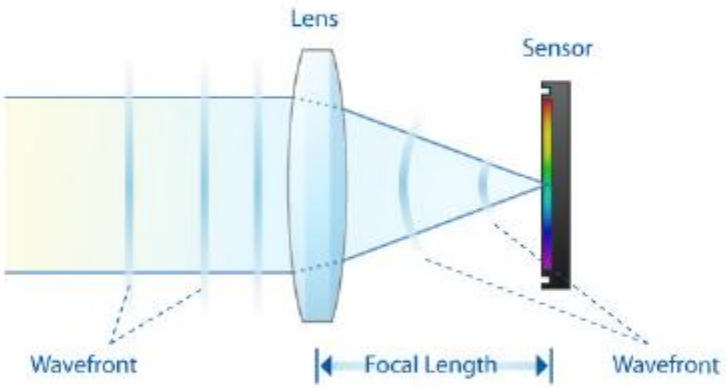
(In astrophotography: sensors are often used in read noise regime, and “dark frames” are used to calibrate the signal level more precisely.)

Comparing raw camera performance

Measuring **sharpness**

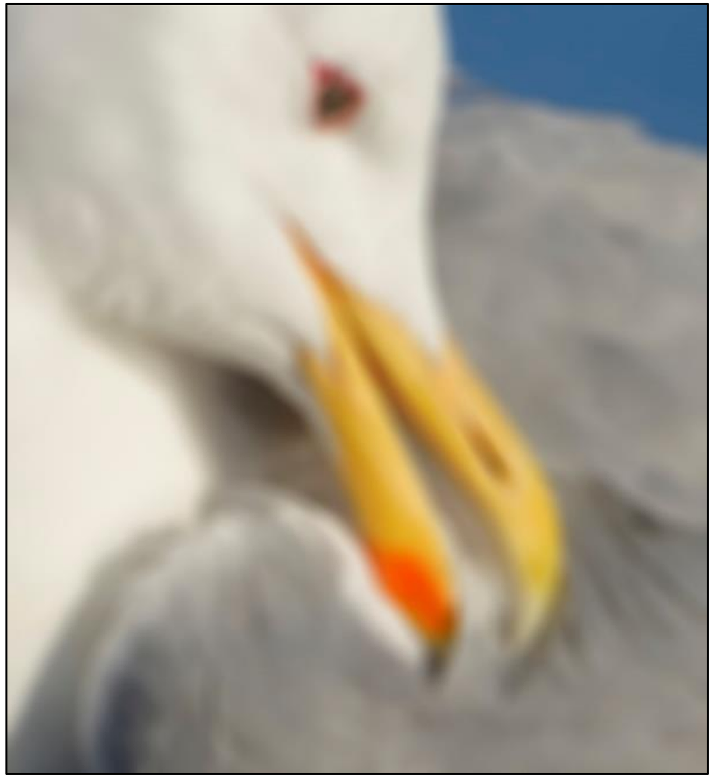
Optical degradation

- The lens introduces degradations to RAW images

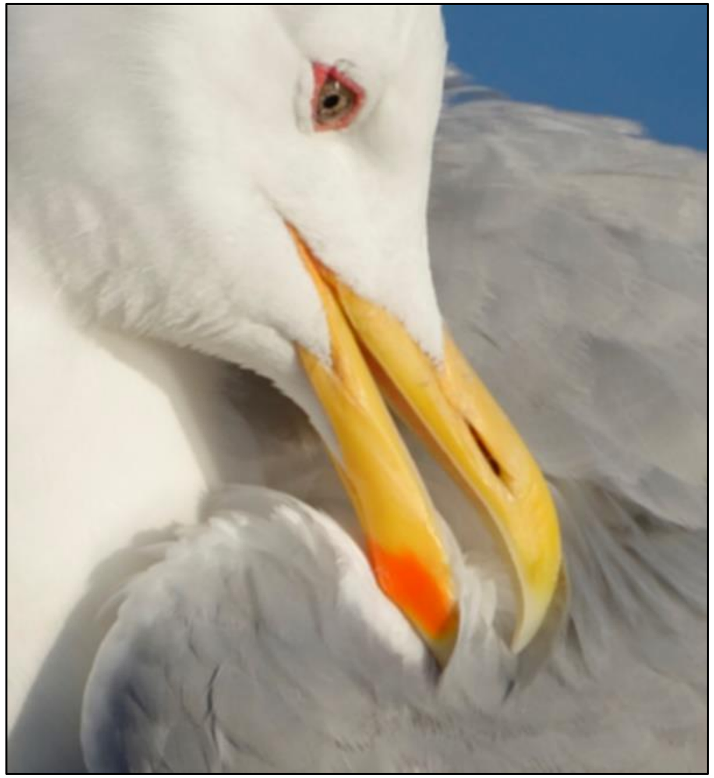


Sharpness

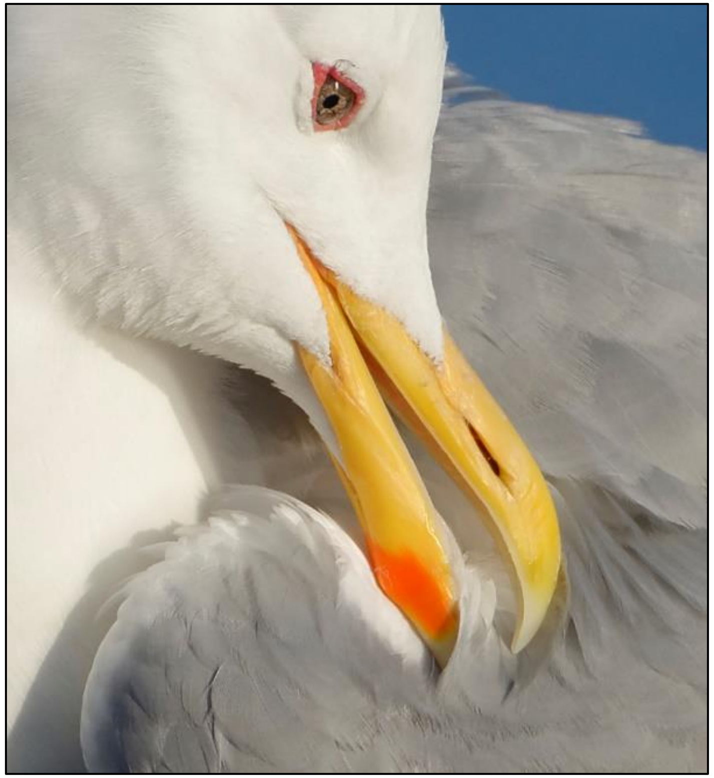
- Different levels of sharpness



Low



acceptable

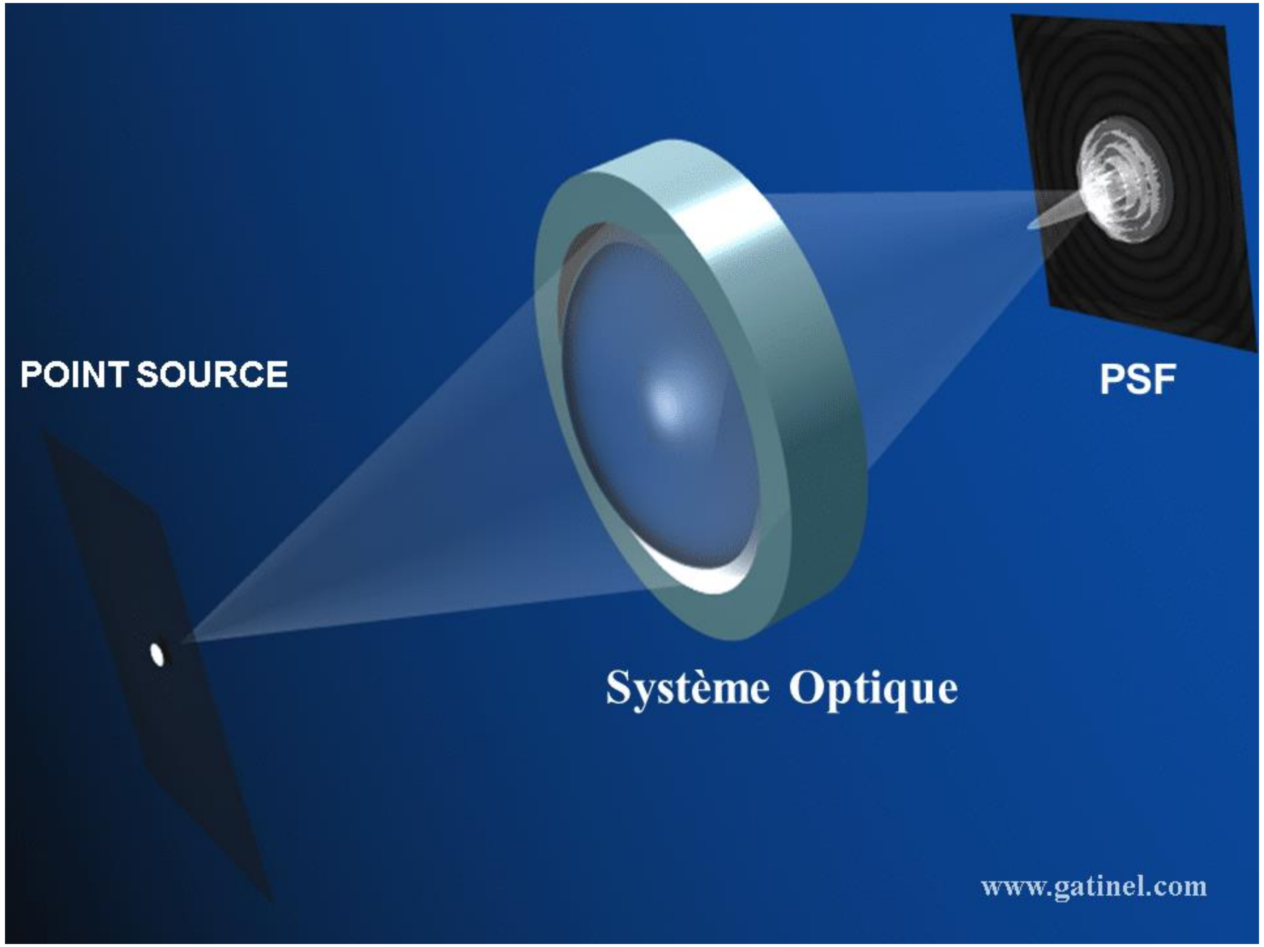


Excellent

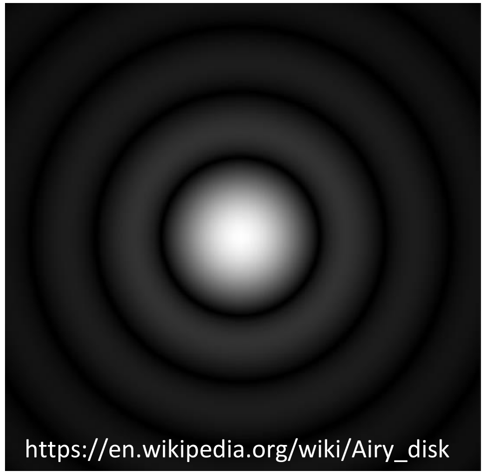
Photos: Frédéric Bourges

Sharpness

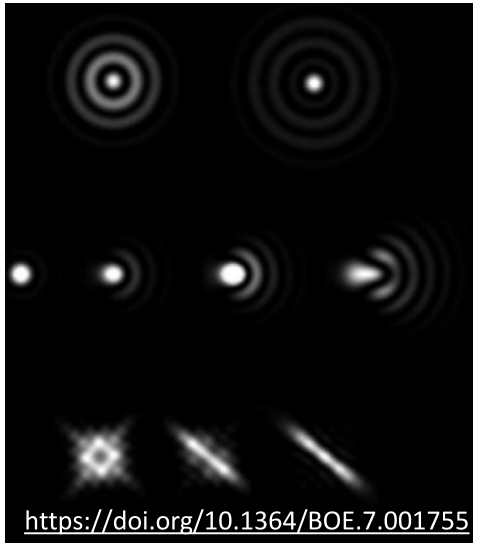
- The image of a point formed by a lens is not a point, but a spot



Sharpness



In a perfect system, the dimension of the spot is due to the diffraction
For a circular aperture, this spot is called Airy disk



For a real system, this spot is called Point Spread Function (PSF)
Size and shape of this spot depend on the aberrations of the lens

Images source : Coles *et al.*, Characterisation of the effects of optical aberrations in single molecule techniques, “Biomedical Optics Express”, 2016

Sharpness

The image of an object formed by a lens is expressed by its convolution with the PSF:

$$\mathrm{Image} = \mathrm{Object} \otimes \mathrm{PSF}$$

The image is less sharp than the object in the scene

In the frequency domain:

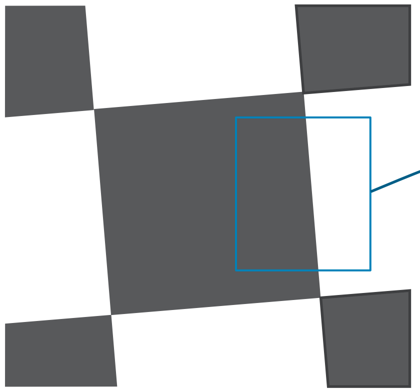
$$\widehat{\mathrm{Image}} = \widehat{\mathrm{Object}} \cdot \widehat{\mathrm{PSF}}$$

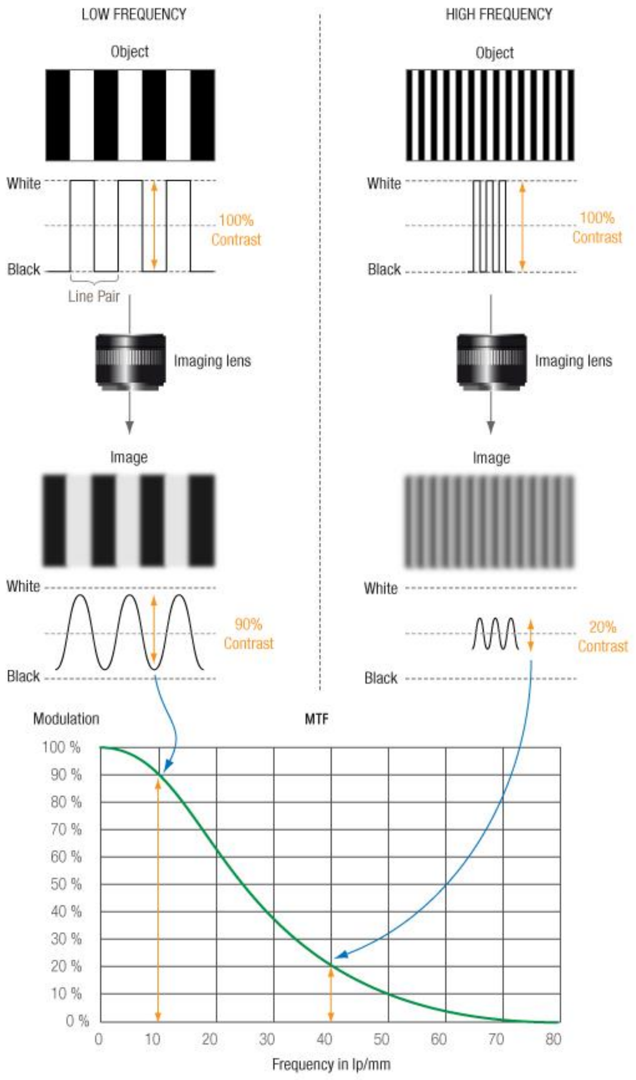
Sharpness – Characterization

In the frequency domain (frequency by frequency analysis), the sharpness is characterized by the Modulation Transfer Function (MTF)

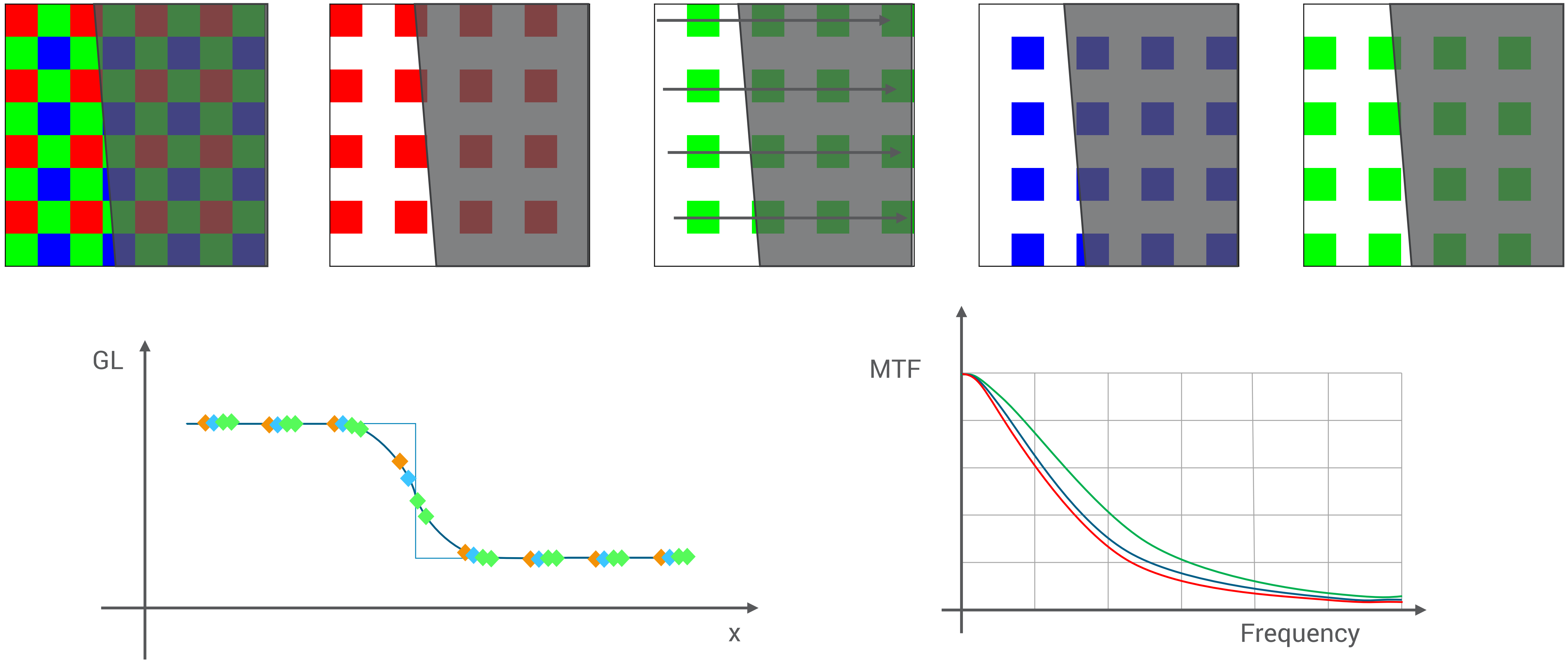
$$\mathrm{MTF} = \left\|\widehat{\mathrm{PSF}}\right\|$$

The MTF represents the attenuation of signal in the image produced by a camera system for each spatial frequency





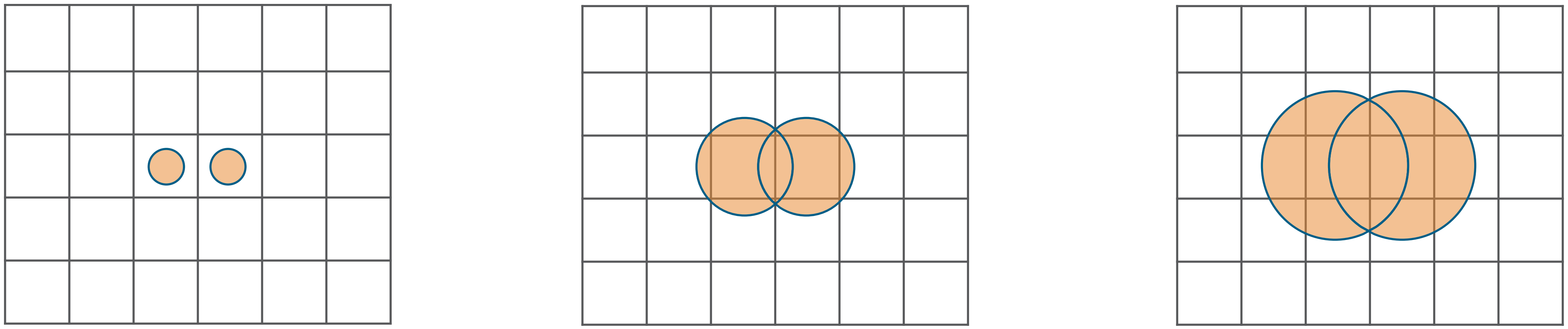
RAW pixel sampling and MTF



Sharpness – Resolution

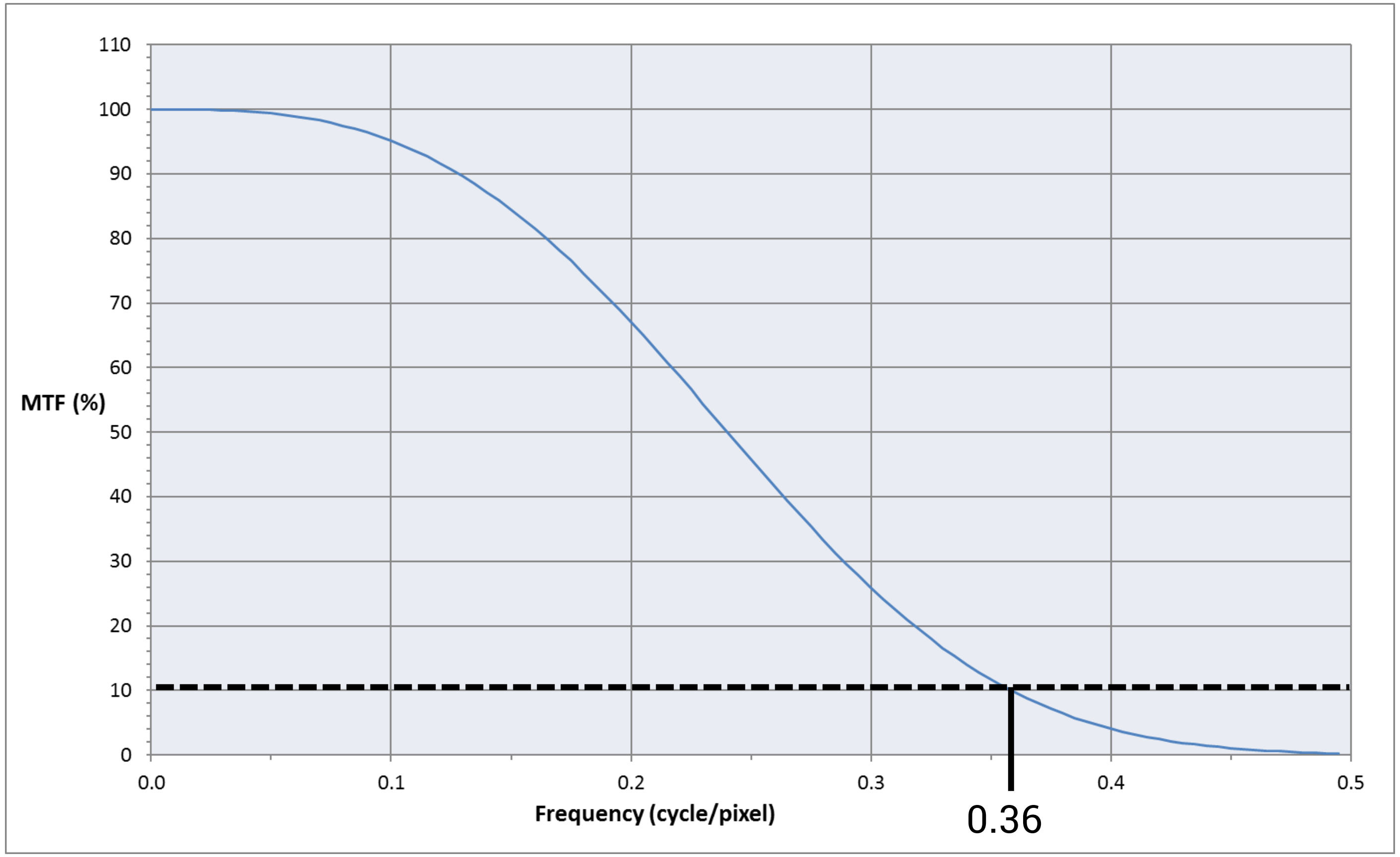
The resolution is the ability to distinguish two closely spaced objects
The resolution depends on:

- the number of pixels (sensor definition)
- the radius of the PSF (optical resolution)



Sharpness characterization

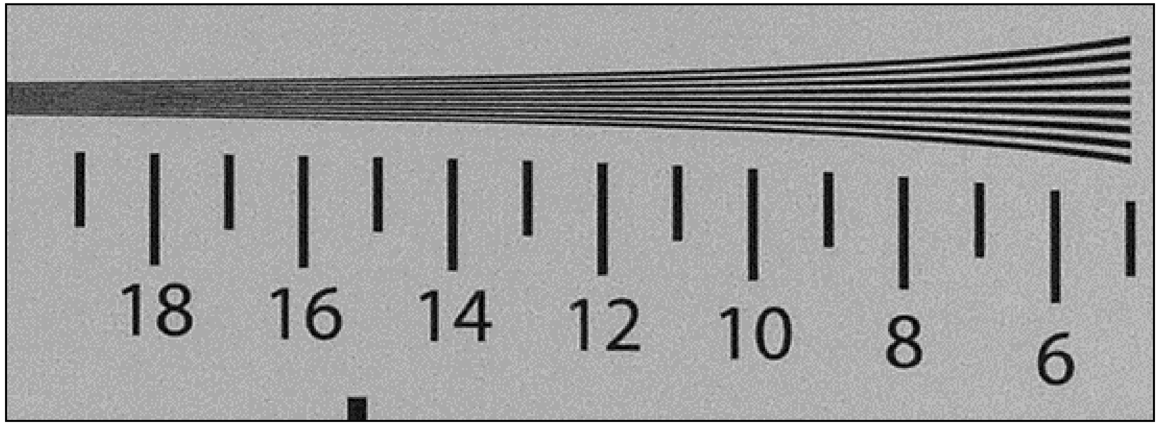
How to extract sharpness information from MTF curves ?

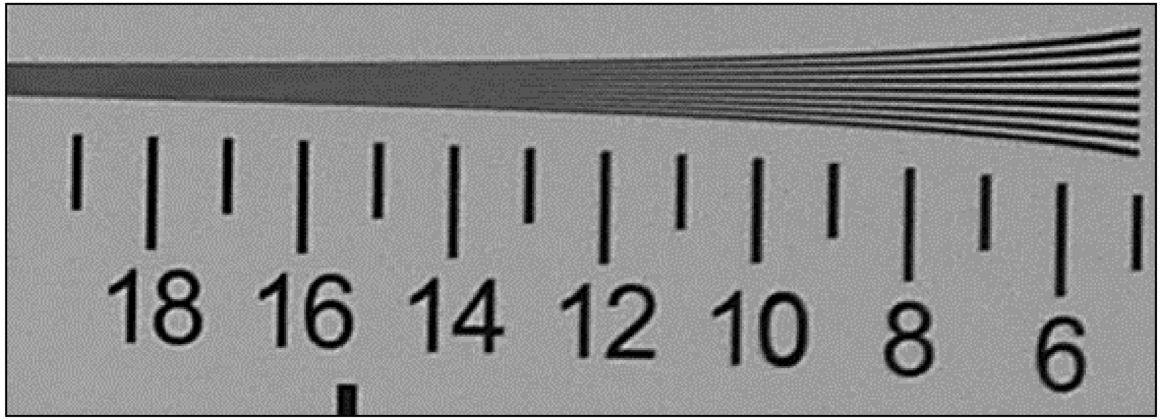


MTF10: 0.36 cy/p
(Limiting resolution)

Sharpness – Resolution

Resolution is related to high frequencies (fine details)

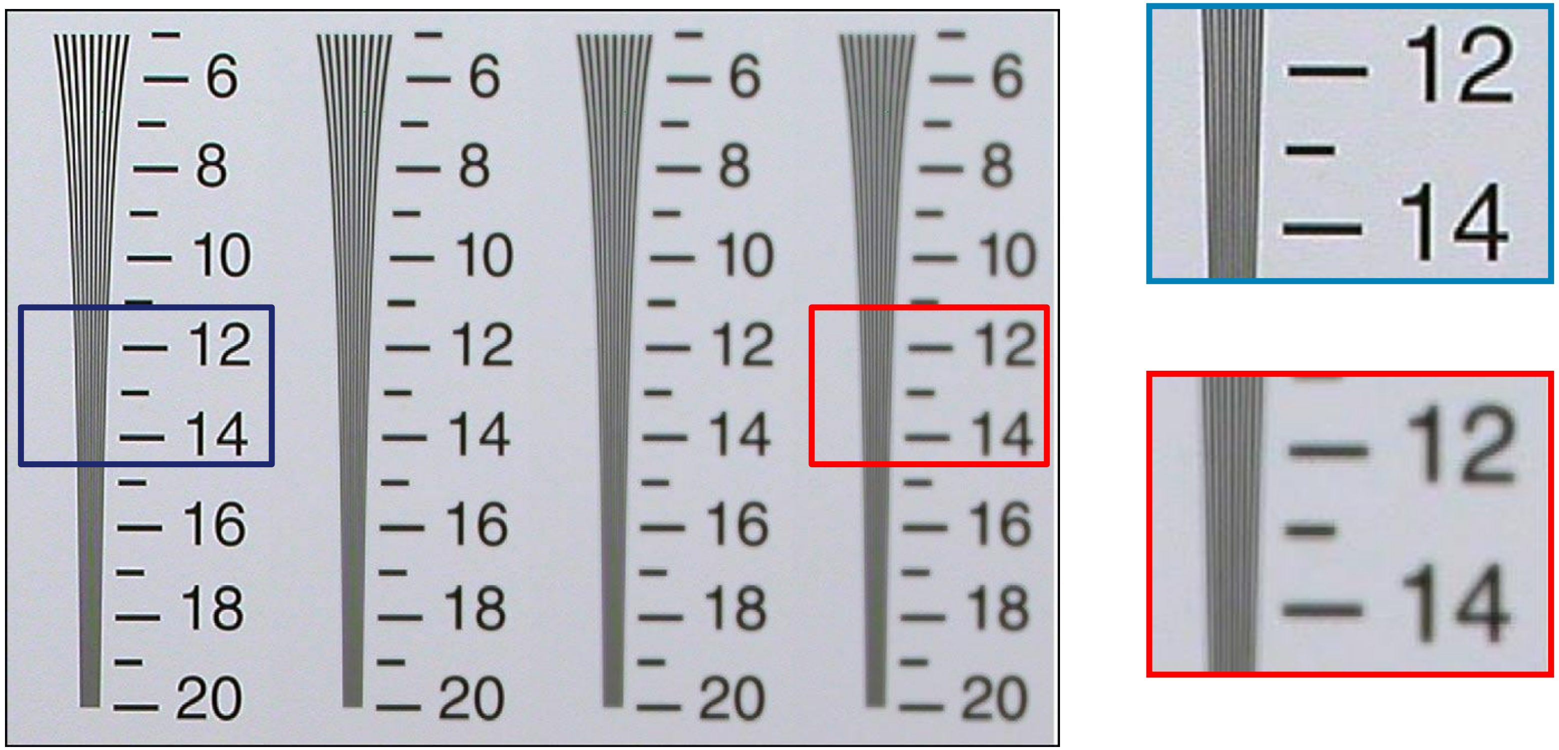




Sharpness – Perception

Perception

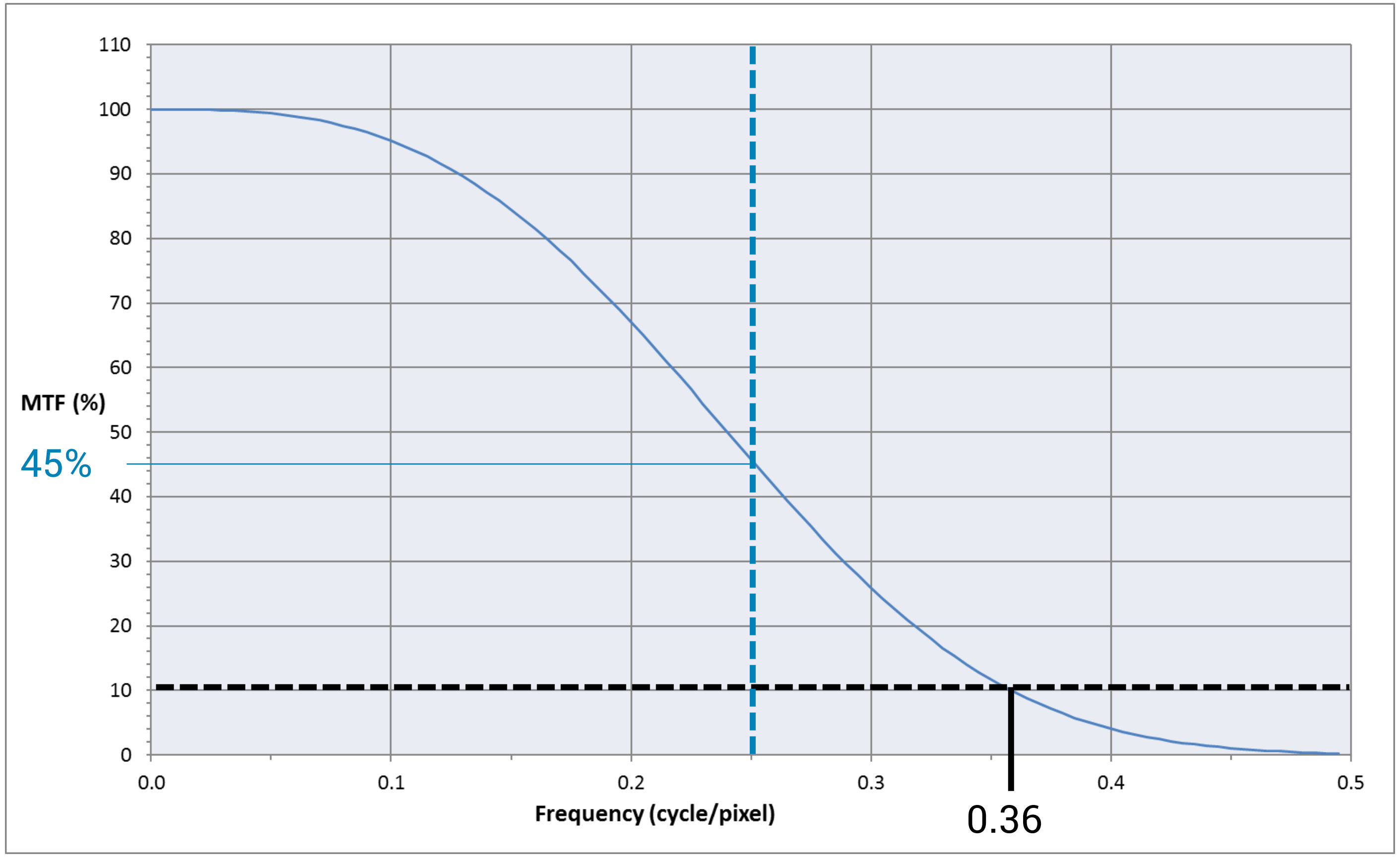
Resolution (small details) ≠ perceptual blur!



Same limiting resolution but increasing perceptual blur level

Sharpness characterization

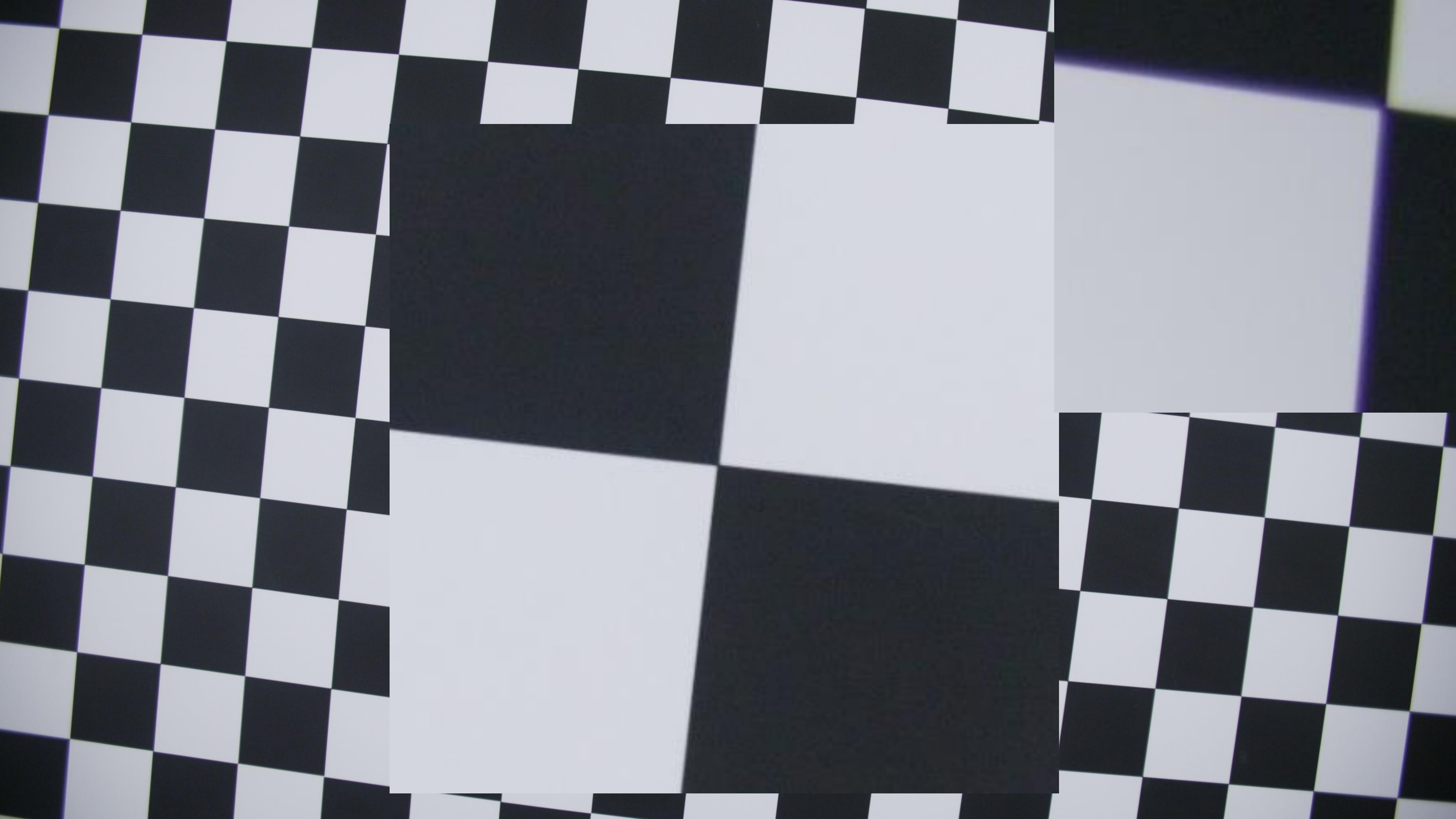
How to extract sharpness information from MTF curves ?



MTF10: 0.36 cy/p

MTF « Half Nyquist » : 45%

Both are very common aggregate metrics for lens characterization!



Comparing raw camera performance

Measuring color response

Color space representation

- In the human eye, there are three types of light sensitive cones, characterized by their spectral response: **S** (for short length), **M** (for medium) and **L** (for long), centered respectively in blue, green, and red.
- The usual color spaces are described by three components. Main hypothesis: a **finite dimensional space can represent all perceived colors**. Color space and spectral space are bijective if and only if there is only one illuminant with no indirect light paths.
- An imaging device can reproduce a portion of that called its **gamut**.
- Other important hypothesis: blending perceived colors is a **linear operation** (in a **linear colorspace**).
- The CIE XYZ **linear color space** (1931) was the first successful modeling based on human color perception and is the *de facto* **common colorspace** used in a large amount of works related to color.
- Other common (not linear) color spaces are: CIELAB (1976), sRGB (1996), Adobe RGB (1998), etc.

Color space representation

Some tips:

CIE-XYZ

Y'CbCr/
Y'UV

- Y is the (linear) **luminance** (in cd/m²)
- ⚠ Y' is the nonlinear encoding or **luma** (unitless)

- Often represented in normalized CIE-xyY **chromaticity coordinates** with:

$$x = \frac{X}{X+Y+Z}, y = \frac{Y}{X+Y+Z}$$

- Monochromatic visible wavelengths form the classic "horseshoe" shape.

- A 3 colors RGB display can output all xy coordinates within a triangle.

Monochromatic wavelengths

Gamut

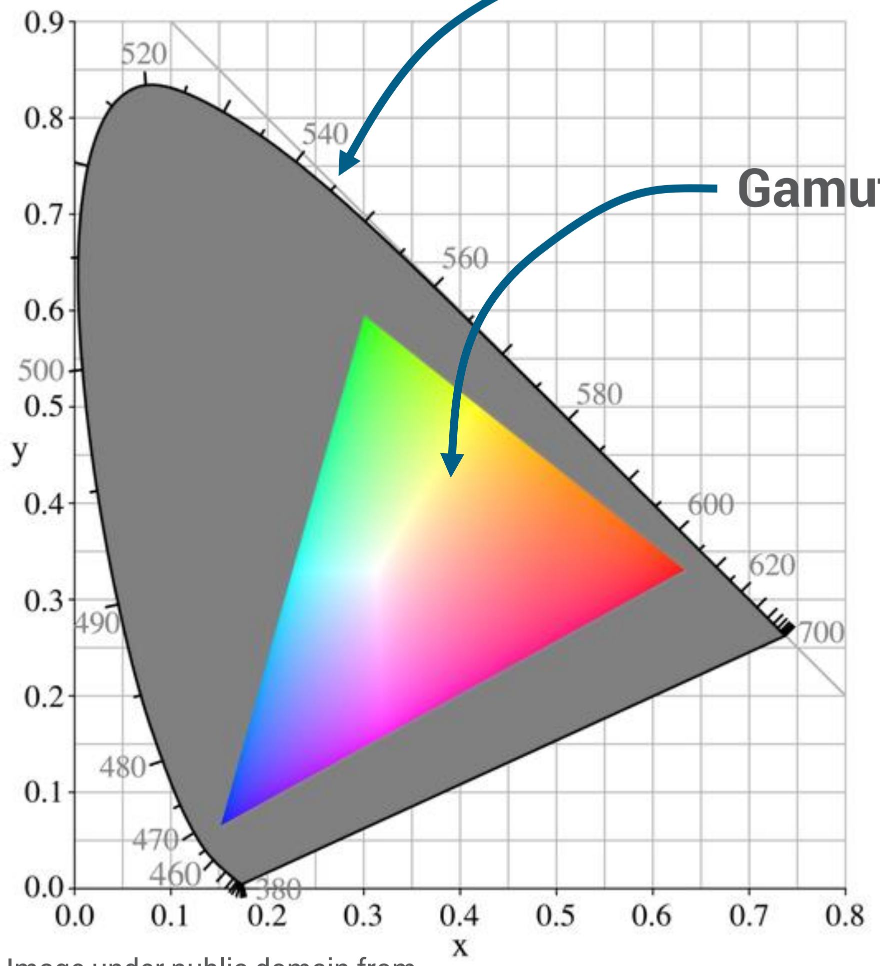
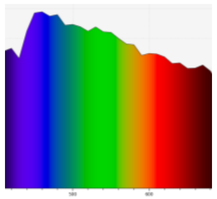


Image under public domain from
https://commons.wikimedia.org/wiki/File:CIExy1931_srgb_gamut.png

Capturing colors with digital cameras

CMOS sensors are (mostly) linear: RGB_{raw} is a linear colorspace!



R_{raw}

G_{raw}

B_{raw}

$$\Rightarrow M_{RAW \to RGB} \Rightarrow$$

$R_{display}$

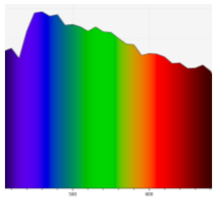
$G_{display}$

$B_{display}$

?

Capturing colors with digital cameras

CMOS sensors are (mostly) linear: RGB_{raw} is a linear colorspace!



R_{raw}

G_{raw}

B_{raw}

$M_{RAW \to sRGB}$

$R_{sRGB\ linear}$ → R'_{sRGB}

$G_{sRGB\ linear}$ → G'_{sRGB}

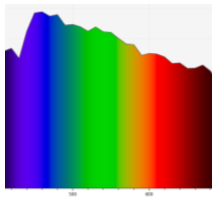
$B_{sRGB\ linear}$ → B'_{sRGB}

Displays are standardized to linear colorspace + a tone curve (often sRGB)

Capturing colors with digital cameras

CMOS sensors are (mostly) linear: RGB_{raw} is a linear colorspace!

Simple multiplicative white balance factors are often useful for algorithmic purposes



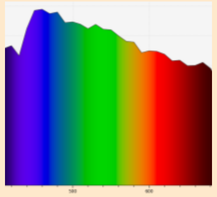
R_{raw}, G_{raw}, B_{raw} → w_{WB} → $M_{WB \to sRGB}$ → $R_{sRGB\ linear}$, $G_{sRGB\ linear}$, $B_{sRGB\ linear}$ → R'_{sRGB}, G'_{sRGB}, B'_{sRGB}

Displays are standardized to linear colorspace + a tone curve (often sRGB)

Note: A lot of cameras also use XYZ conversion matrices $M_{WB \to sRGB} = M_{XYZ \to sRGB} \cdot M_{WB \to XYZ}$

Capturing colors with digital cameras

Illuminant spectrum: CIE-A



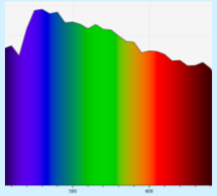
R_{raw}, G_{raw}, B_{raw} → w_A → $M_{A \to sRGB}$ → $R_{sRGB\ linear}$, $G_{sRGB\ linear}$, $B_{sRGB\ linear}$

$R_{sRGB\ linear}$ → R'_{sRGB}

$G_{sRGB\ linear}$ → G'_{sRGB}

$B_{sRGB\ linear}$ → B'_{sRGB}

Illuminant spectrum: CIE-D65



R_{raw}, G_{raw}, B_{raw} → w_{D65} → $M_{D65 \to sRGB}$ → $R_{sRGB\ linear}$, $G_{sRGB\ linear}$, $B_{sRGB\ linear}$

$R_{sRGB\ linear}$ → R'_{sRGB}

$G_{sRGB\ linear}$ → G'_{sRGB}

$B_{sRGB\ linear}$ → B'_{sRGB}

...

Capturing colors with digital cameras

Example characterization for one illuminant:

- White balance factors w_{WB}
- Color matrix $M_{WB \to sRGB}$

White balance scales	R_{RAW}	2.4
	G_{RAW}	1.0
	B_{RAW}	1.6

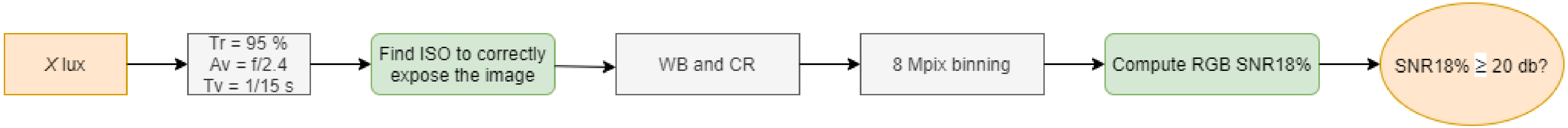
		R_{sRGB}	G_{sRGB}	B_{sRGB}
Color matrix as defined in ISO standard 17321	R_{RAW}	1.72	-0.49	-0.24
	G_{RAW}	-0.21	1.51	-0.30
	B_{RAW}	0.00	-0.57	1.57

Measuring ability to capture colors with digital cameras

What makes a sensor have good/bad color rendering by itself?

- Impact of white balance / color matrix on noise:
  - SNR10: Minimum illuminance under which RGB noise is too degraded
  - Color Sensitivity (CS): Number of distinguishable RGB colors, up to noise
- Impact of nonlinearities on color rendering accuracy:
  - Close reproduction of XYZ sensitivities (metamerism, SMI: Sensitivity Metamerism Index)
  - Sensor response nonlinearities (color rendering is different in shadows/highlights)
- Perception of colors?

The SNR10 is smallest lux level such that the corresponding ISO speed leads to an RGB SNR18% greater or equal to 10 (i.e., 20 dB):

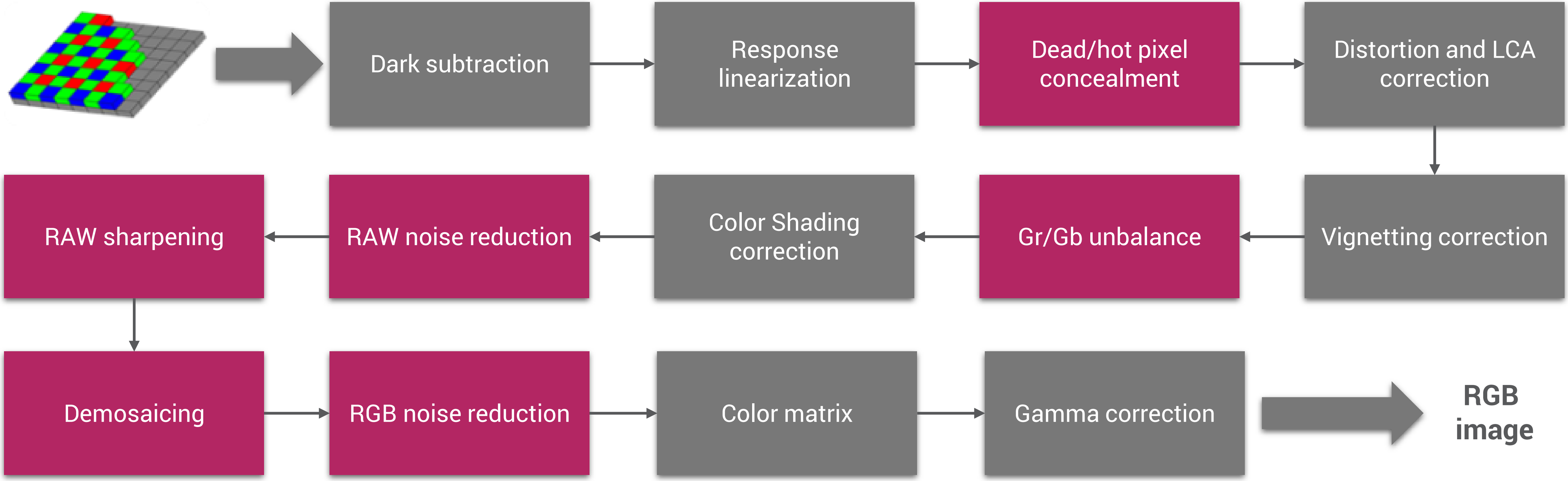


Comparing cameras for photography (JPEG output)

A short recap on RGB/JPEG formats

ISP: A complex image pipeline from RAW to RGB

- A typical pipeline includes several steps that are non-linear and non-invertible



Comparing cameras for photography (JPEG output)

Measuring color perception & fidelity

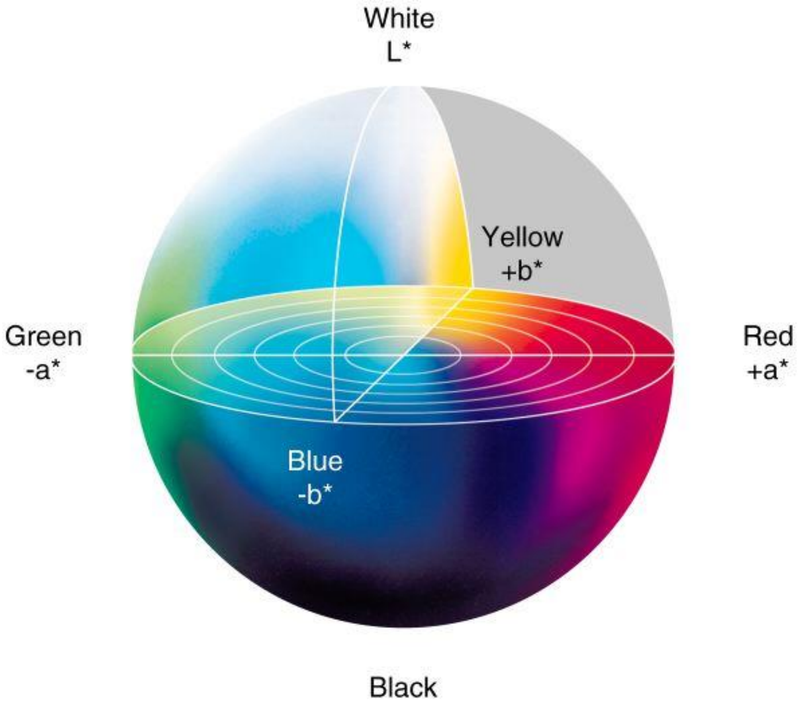
Photo V5 protocol | **Color** – How to evaluate Color Fidelity?

To assess Color Fidelity, it is required to measure color difference (or distance).

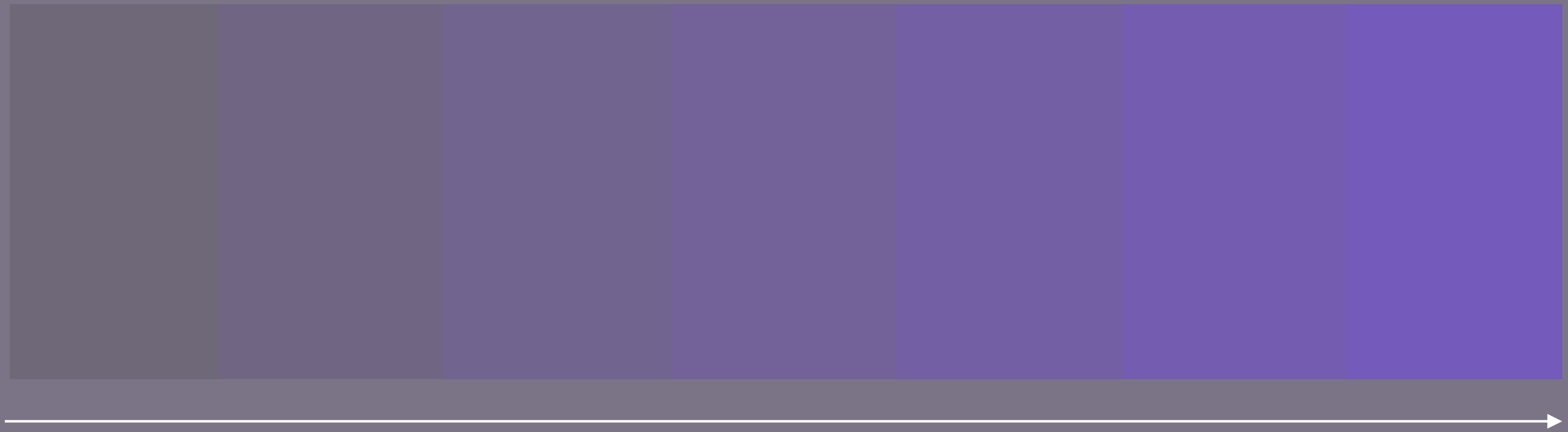
Problem : The sRGB color space is not perceptually uniform under the Euclidean norm (*in fact, no usual norm has this property in the sRGB space*).

This need leads to the **CIELAB** (also called **L*a*b***) color space, created in 1976.

- This space is an **opponent color space** that separates lightness (L^*) and chroma (a^*, b^*).
- Approximately **perceptually uniform** under the Euclidean norm.

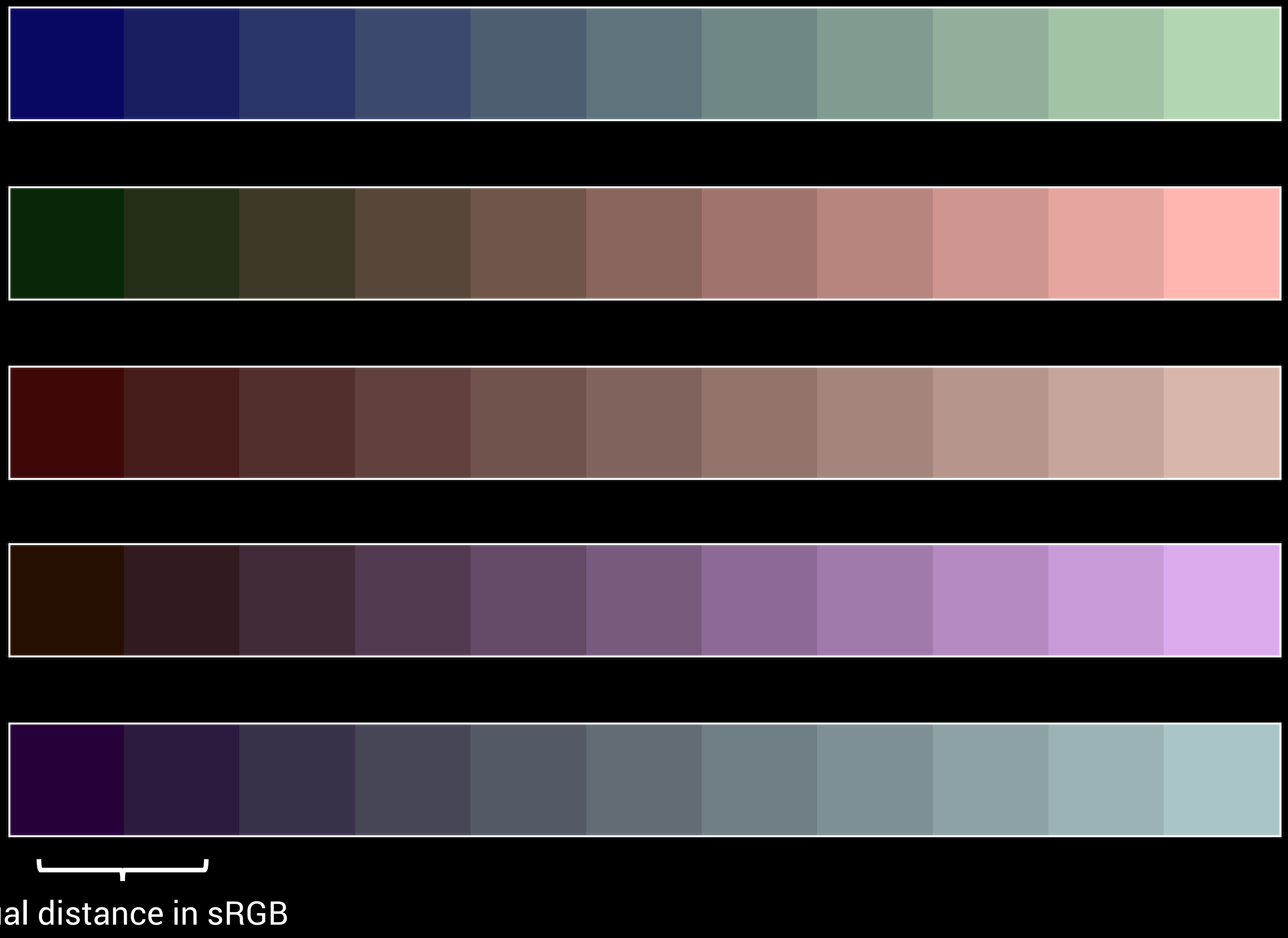


Increasing L^* does not change colors

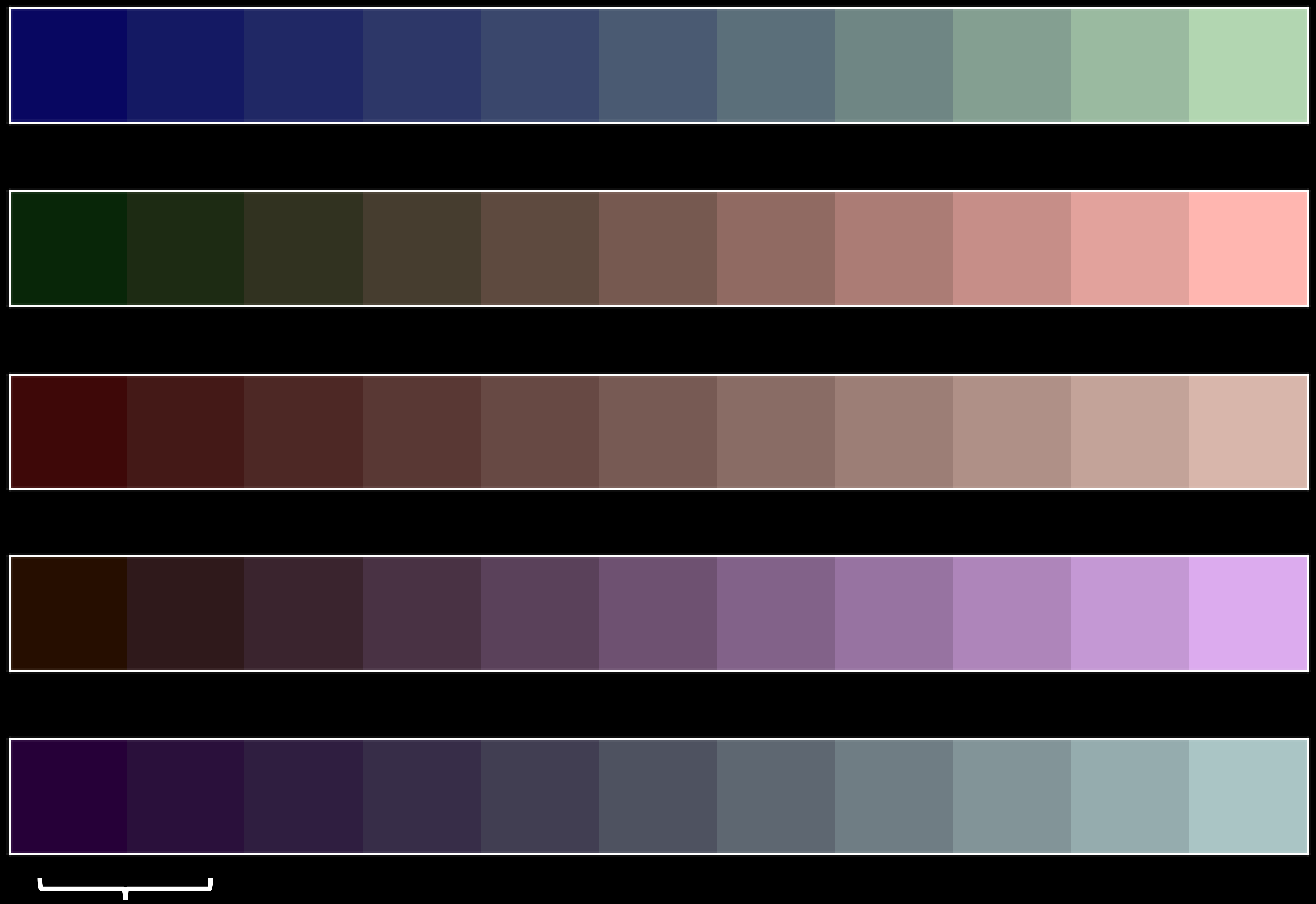


Increasing a^*, b^* does not change lightness

sRGB



CIELAB



Equal distance in CIELAB

Photo V5 protocol | **Color** – How to evaluate Color Fidelity?

Distances between colors:

The Euclidian distance CIE 1976:

$$\Delta \boldsymbol{E}^* = \sqrt{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}},$$

$\Delta E^* > 3$ represents a noticeable color difference.

Absolute Chroma error:

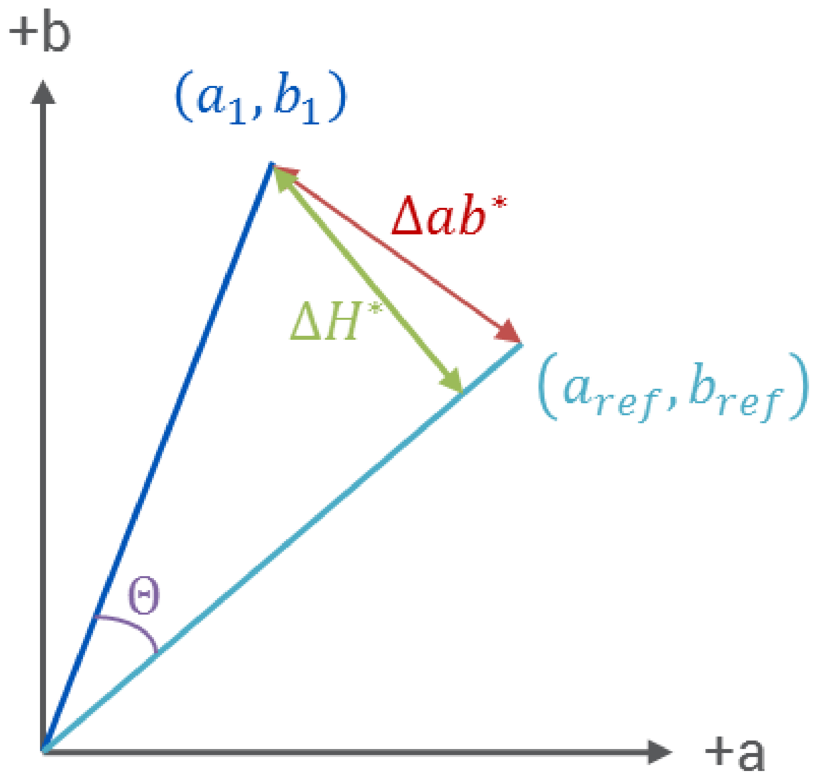
$$|\boldsymbol{\Delta C}^*| = \left| \sqrt{a_1^{*2} + b_1^{*2}} - \sqrt{a_{ref}^{*2} + b_{ref}^{*2}} \right|$$

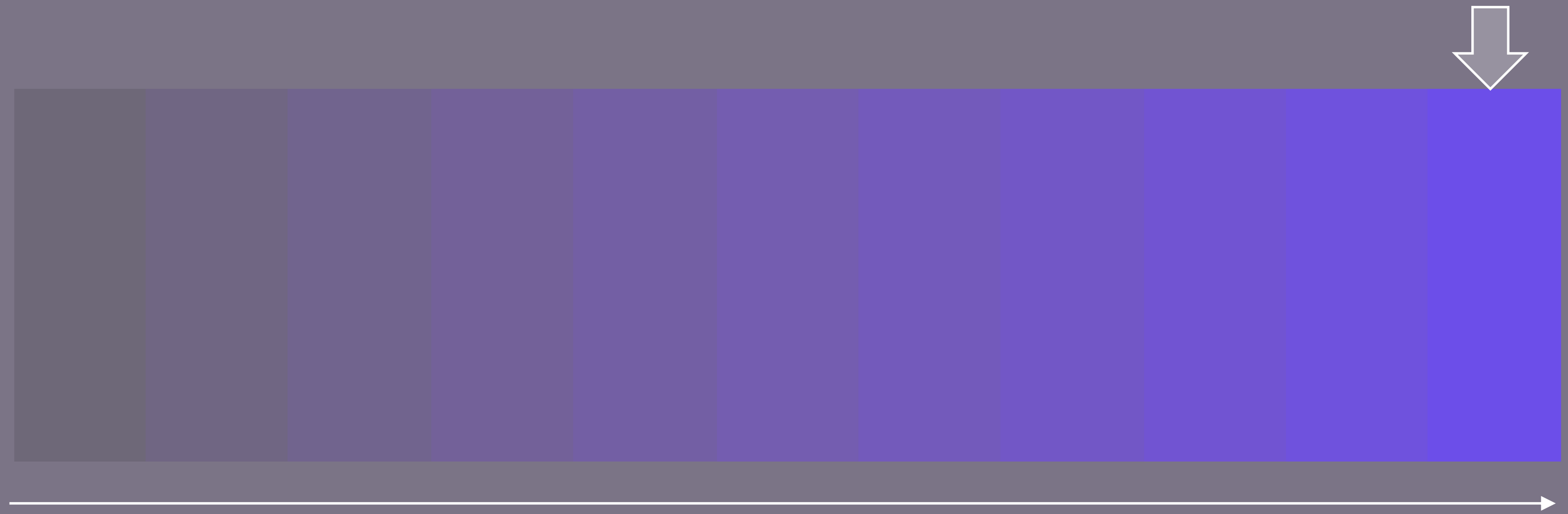
Hue error:

$$\boldsymbol{\Delta H}^* = \sqrt{\Delta E^{*2} - \Delta L^{*2} - \Delta C^{*2}}$$

Color difference (without exposure):

$$\boldsymbol{\Delta ab}^* = \sqrt{\Delta a^{*2} + \Delta b^{*2}}$$





Increasing a^*, b^* does not change lightness
…or does it?

Photo V5 protocol | **Color** – Color fidelity limits

All **perceptual color models have inherent limitations** and are generally optimized for a particular purpose.

While Δab^* gives indication on the fidelity, **it is not related to people preferences**, who tend to prefer more saturated colors and warm casts.

Color Fidelity evaluation alone is not sufficient to reflect the user experience!

These limits lead to split the color accuracy evaluation into 2 parts:

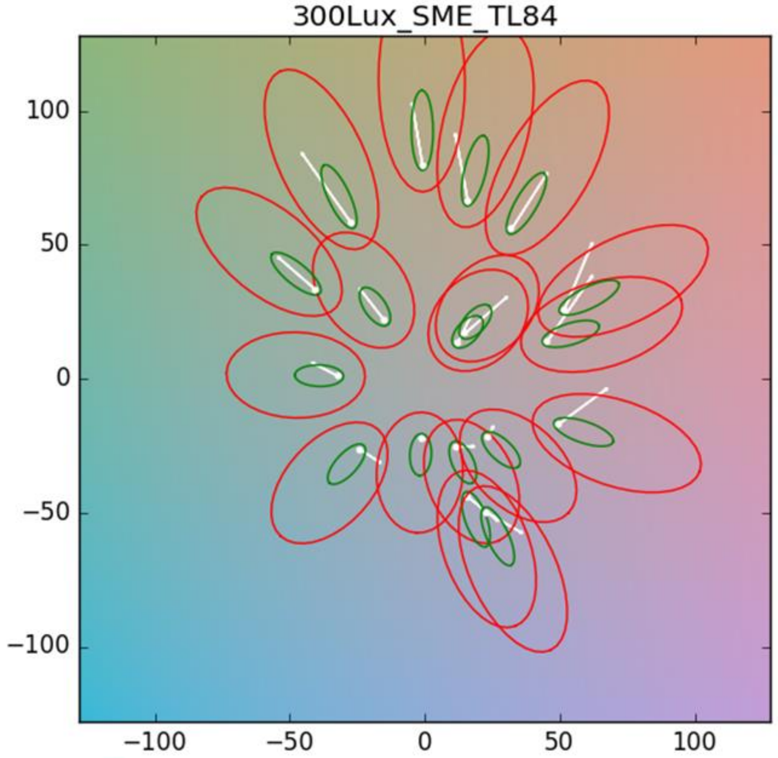
- White balance:
  - Yellow and orange casts, are preferable to green, pink and blue casts.
- Color Rendering:
  - Accuracy in terms of hue is more important than saturation accuracy.

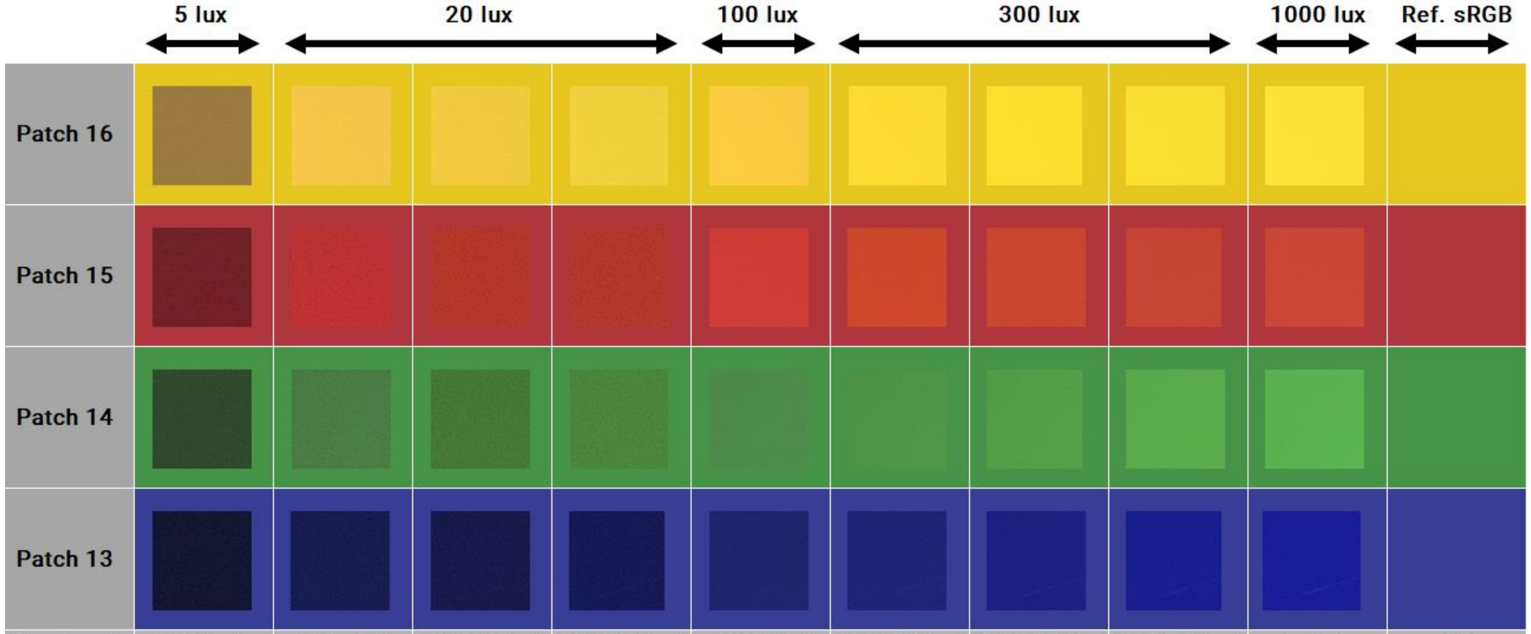
Photo V5 protocol | **Color measurements** – Color Rendering

To evaluate the hue and acceptable level of color saturation, we designed ellipses for each patch of the ColorChecker®:

The green ellipses indicate the boundaries of the best score.

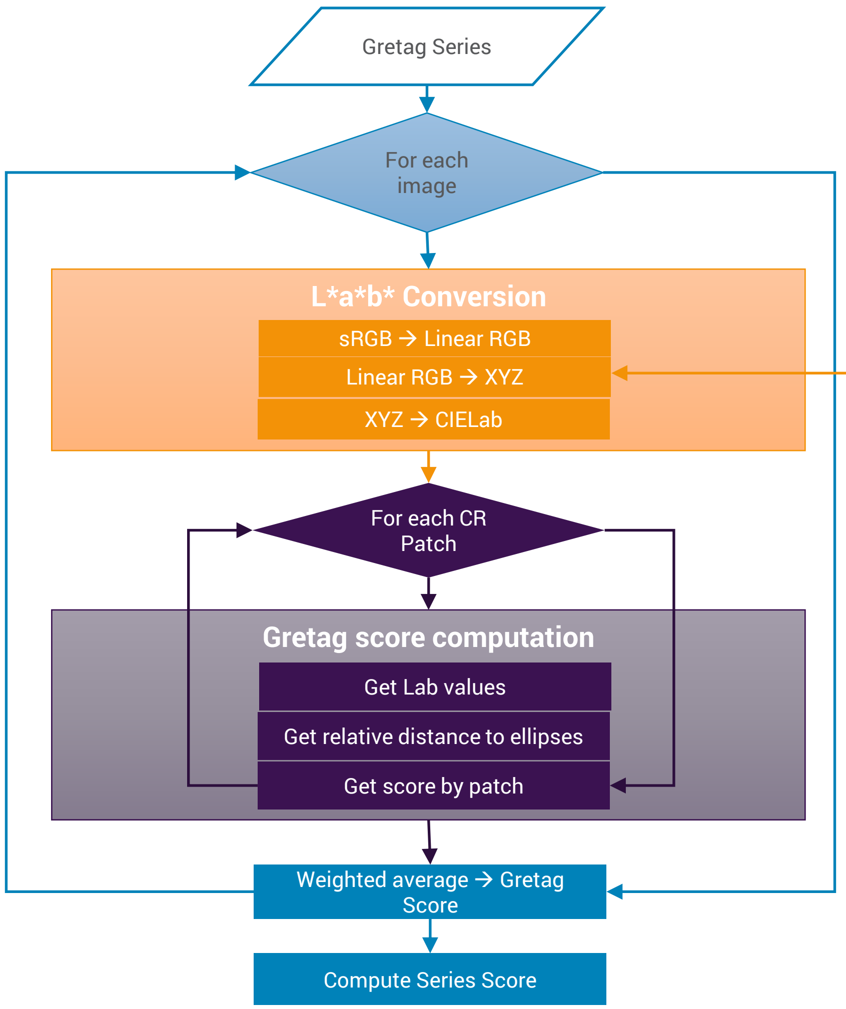
The red ellipses indicate the boundaries of the worst score.





Illuminant	H (5 lux)	A (20 lux)	TL83 (20 lux)	TL84 (20 lux)	A (100 lux)	TL83 (300 lux)	TL84 (300 lux)	D65 (300 lux)	Daylight	D65
Delta ab	27.5	12.7	12.6	11.1	11.4	17.9	21.3	22.6	24.1	
Delta C	27.0	10.3	10.7	9.8	9.4	15.6	19.5	21.0	22.6	
Delta H	3.5	4.6	5.5	4.8	5.7	6.2	6.7	6.9	6.9	

Photo V5 protocol | **Color measurements** – Color Rendering



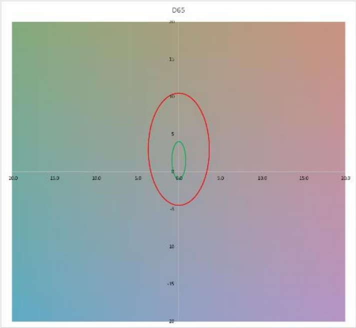
White point properties used for conversion adapted to the target illuminant

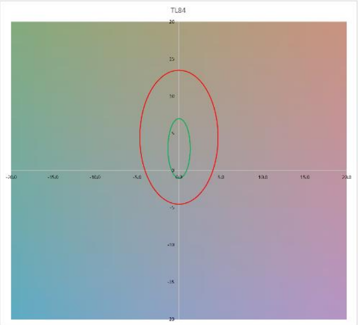
Photo V5 protocol | **Color measurements** – White balance

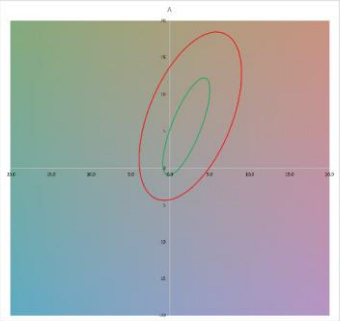
Objective measurements of white balance are made on the gray patches of ColorChecker®.

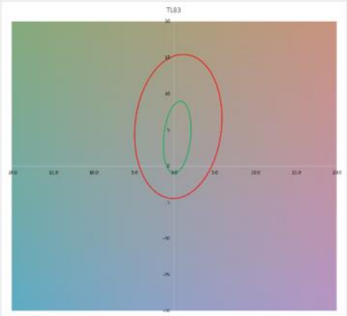
Ellipses are different, depending on the illuminant color temperature.

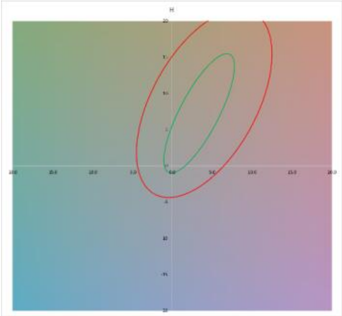
	5 lux	20 lux			100 lux	300 lux			1000 lux	Ref. sRGB
Patch 19										
Patch 21										
Patch 23										
Illuminant	H (5 lux)	A (20 lux)	TL83 (20 lux	TL84 (20 lux)	A (100 lux)	L83 (300 lux	L84 (300 lux	D65 (300 lux)	Daylight	D65
Delta ab	10.9	14.6	17.1	13.1	14.2	12.5	8.9	6.2	3.5	
Delta C	9.5	13.2	15.7	11.8	12.8	11.1	7.5	4.8	2.1	
Delta H	5.3	6.2	6.8	5.9	6.2	5.7	4.8	3.9	2.8	
Hue	69.9	72.4	75.3	84.6	75.4	81.8	87.0	101.6	110.3	
WB repeatabil	0.7	0.4	0.2	0.2	0.8	1.0	0.5	0.2	0.1	





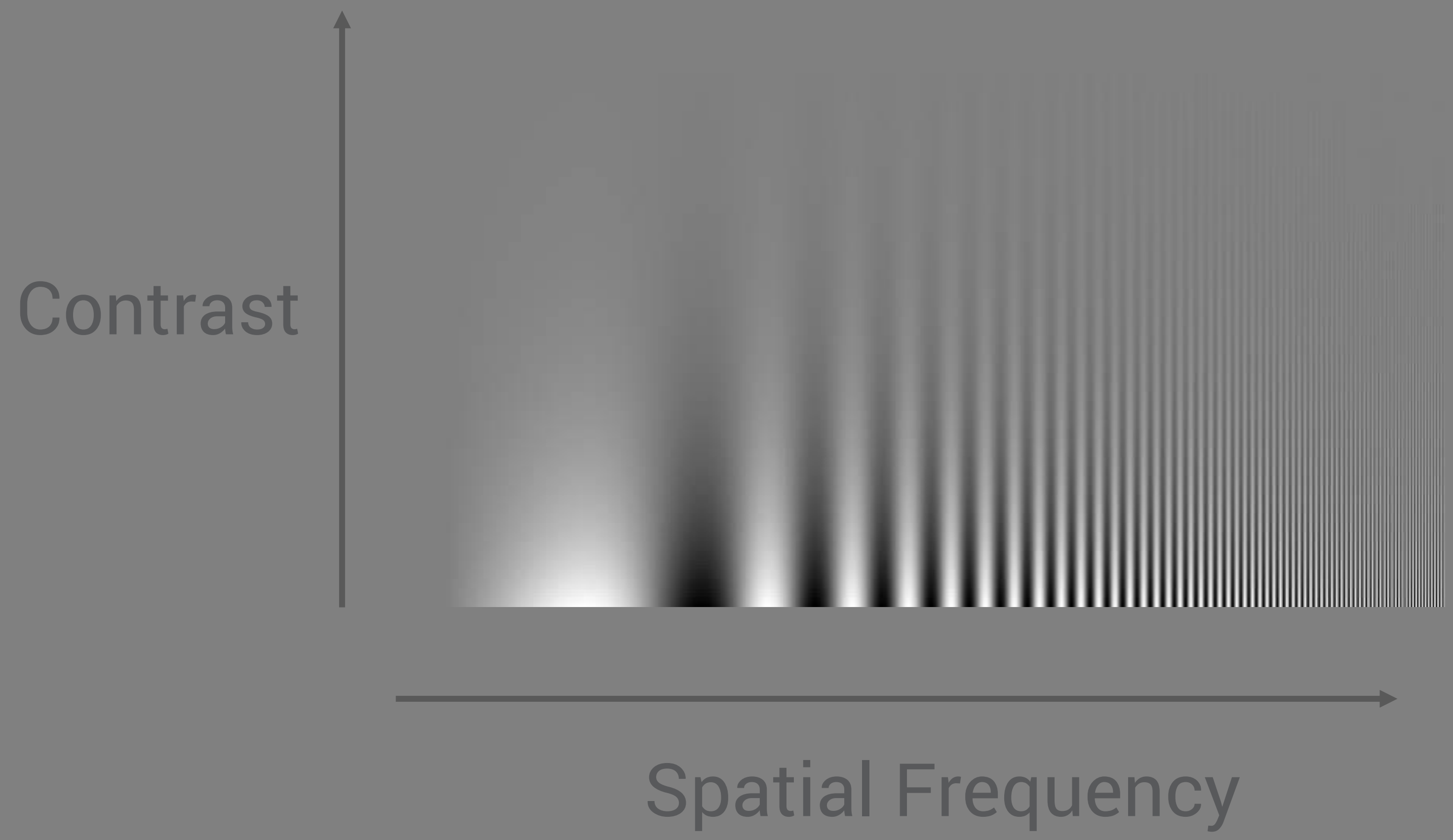


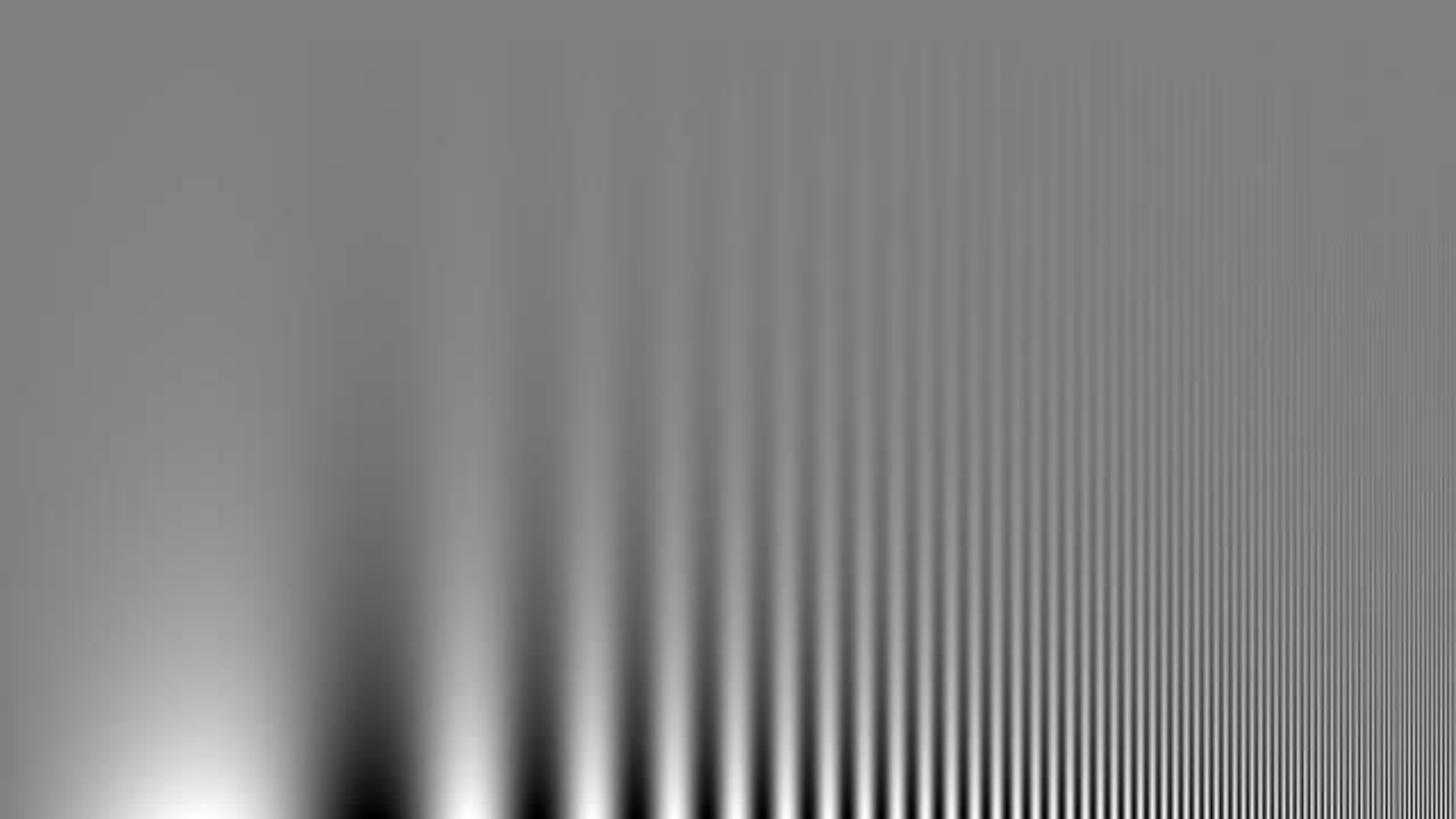




Comparing cameras for photography (JPEG output)

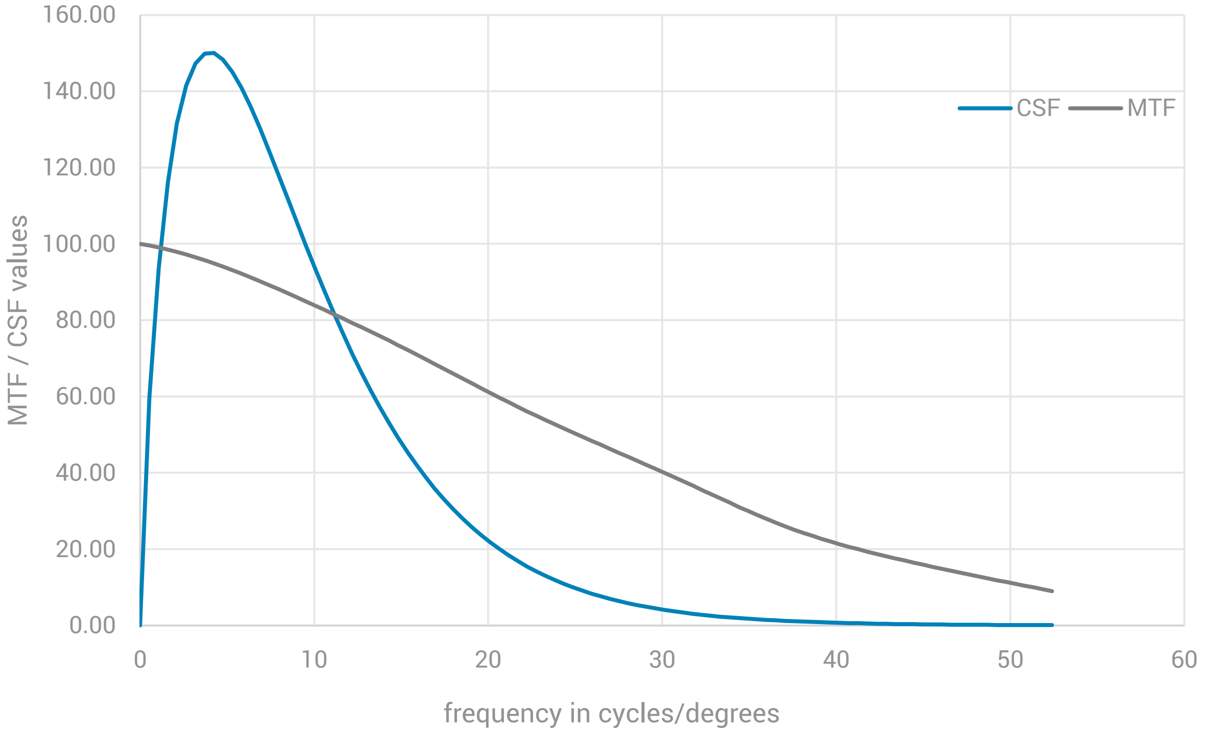
A short introduction to viewing conditions


Contrast
Spatial Frequency



Viewing conditions – CSF (Contrast Sensitivity Function)

Here is the shape of the contrast sensitivity function depending on a selected viewing condition:



CSF** for a 40x60 print at 60cm, alongside a **MTF

Viewing conditions – Acutance

Acutance is a single-value metric calculated from an MTF:

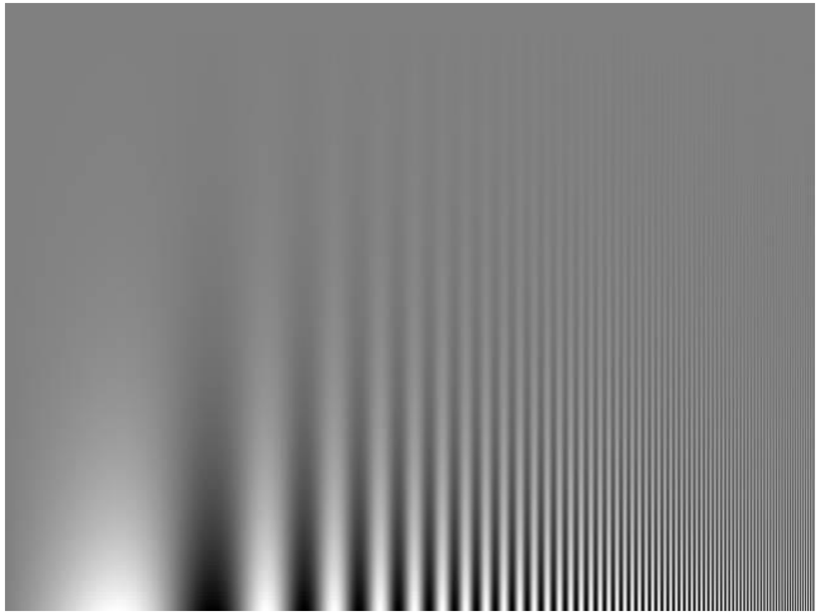
- It uses a **Contrast Sensitivity Function** (CSF), modeling the spatial response of the human visual system to weigh the values of the MTF for the different spatial frequencies. A higher acutance means a sharper image.
- CSF expressed in cycle/pixel depends on the viewing conditions and the image resolution.

Acutance formula is:

$$Acutance = \frac{A}{A_r}$$

with $A = \int_0^\infty MTF(v).CSF(v).dv$,

and $A_r = \int_0^\infty CSF(v).dv$.



Intuitive illustration of the CSF.

Comparing cameras for photography (JPEG output)

Measuring visual noise

Visual noise

- Fine (high-frequency) noise is preferable to coarse (low-frequency).

Fine grain noise

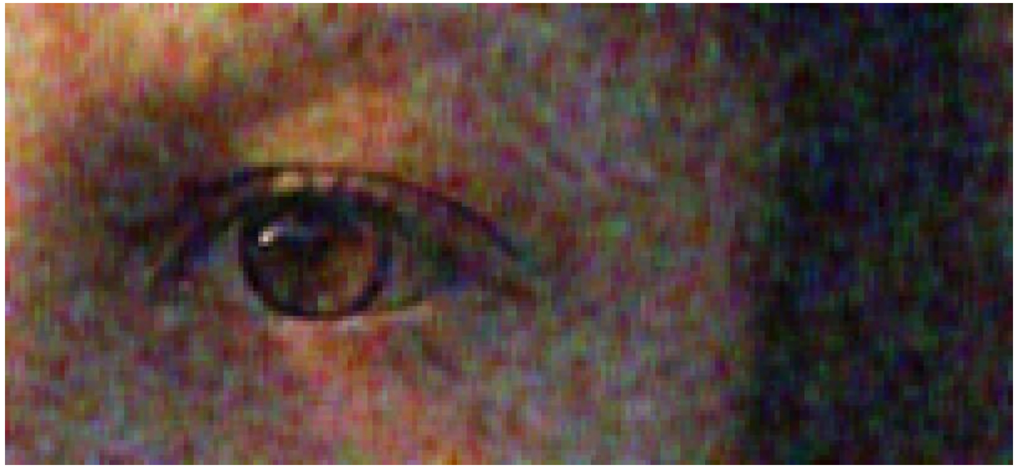
Coarse grain noise

Visual noise

- Luminance noise is preferable to chromatic noise

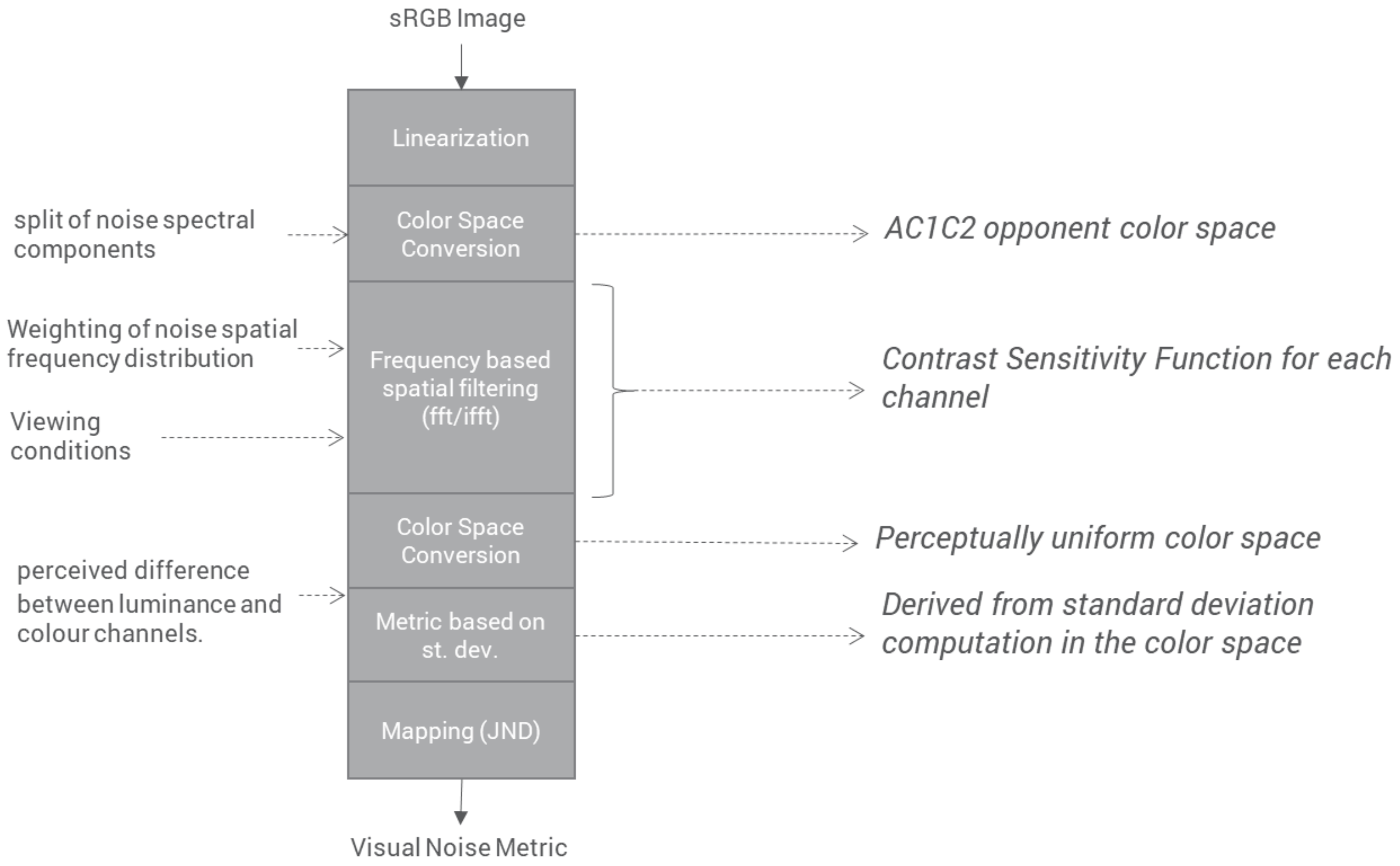
Luminance noise

Chromatic noise



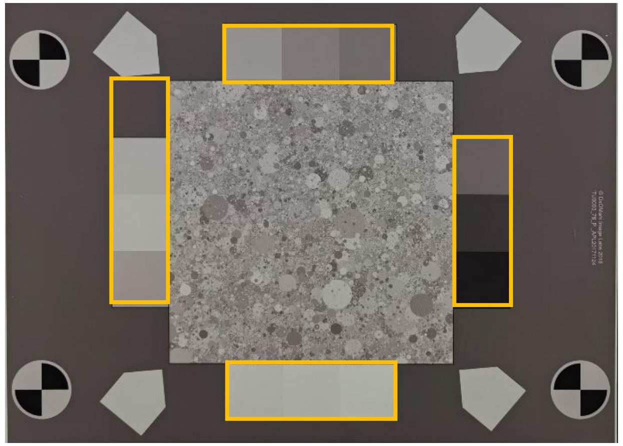
Visual noise

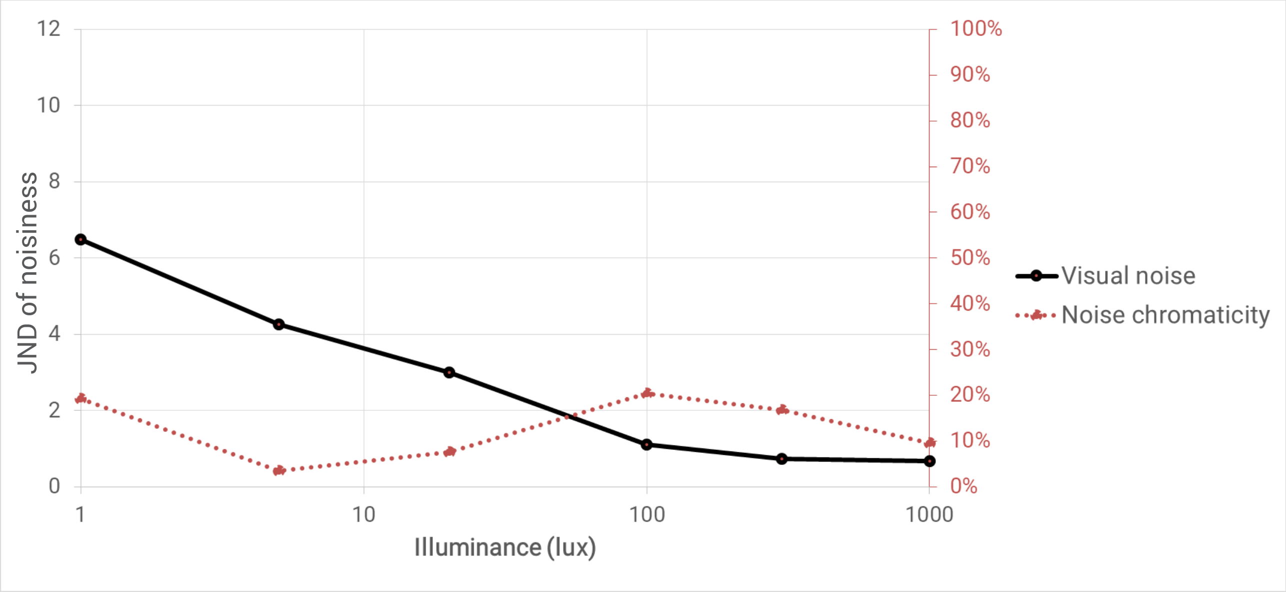
The model weights spectral components of the noise and takes into account the noise spatial frequency distribution, viewing conditions, and the perceived difference between luminance and color channels.



Visual noise

Objective measurements of luminance and chromatic noise are performed on all the gray patches of the Dead Leaves chart.





Non-Linear JND Mapping

Visual noise is displayed with a JND unit scale which is easier to understand perceptually: 75% of people can see a difference of noise when there is 1JND of difference between two patches/images.

Noise chromaticity

Noise chromaticity ratio is the ratio of the weighted chrominance a, b* variance with respect to the weighted chrominance and lightness L* variance (in CIELAB color space). It gives an indication of the chromaticity of the noise but is not used in the score computation.*

Comparing cameras for photography (JPEG output)

Measuring **detail preservation**

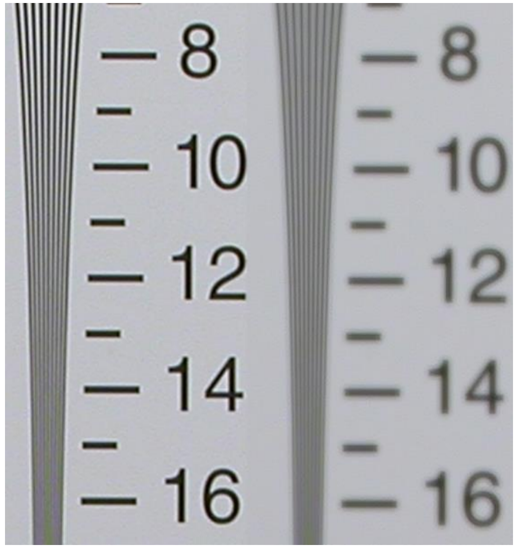
Photo V5 protocol | **Texture & Noise** – Texture preservation

We want to quantify the amount of texture
(i.e. **low contrast fine details**) the camera is able to preserve.

Different amounts of texture

Which differs from

- Resolution: the smallest visible detail perceived by a human observer,
- Sharpness: how acute boundaries of objects are.



Same resolution, different sharpness



Photo V5 protocol | **Texture & Noise** – Texture preservation

Blur and content based digital image processing can reduce the amount of texture in an image.

What causes blur?

- The gap between real lenses and the ideal optical model using thin lenses,
- Focus failures and/or depth of field,
- Motion blur caused either by subject or user,
- Post processing algorithms,
- Limited sensor resolution.

Digital image processing like denoising can remove some of the details wrongly considered as noise.

Photo V5 protocol | **Texture measurements** – MTF

MTF is classically estimated on a **slanted edge** (ISO 12233).

That works well for optical systems with no digital content based post-processing.

Modern camera-phones embed heavy image processing and can perform better on such a simple structure than on natural images with textured areas.

We need a new pattern of **low contrast fine details** to estimate camera's performances of **texture preservation**.

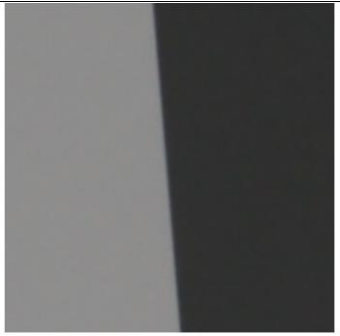
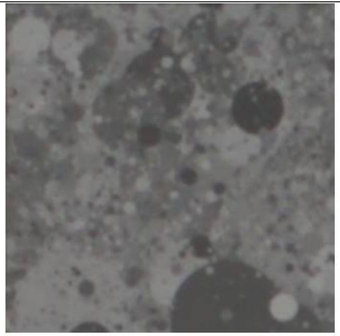
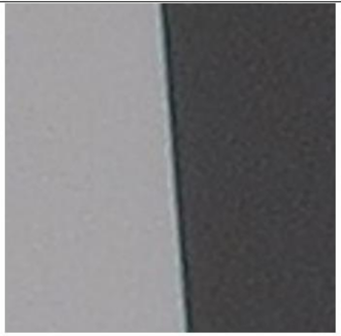
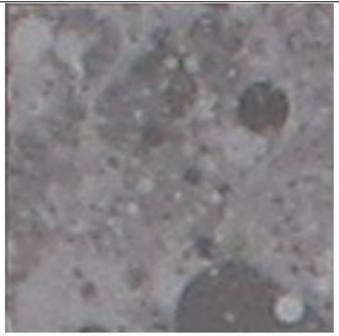
Camera A: DSL-R without edge enhancement		
Camera B: Bridge camera with high edge enhancement		

Photo V5 protocol | **Texture measurements** – Acutance (perceptual)

Texture acutance reflects the detail preservation (low contrast details).
Edge acutance reflects the sharpness (high contrast edges).

On the left, fine details are better preserved. On the right, image is sharper.

Photo V5 protocol | **Texture measurements** – Dead Leaves Pattern

We use a "Dead Leaves" pattern to estimate the acutance. This pattern has several interesting properties:

- Follows the distribution of the statistics in natural images
- Is scale invariant
- Is rotation invariant
- Is shift invariant
- Is contrast/exposure invariant
- Contains edges and occlusions like natural textures
- Is not easily improved by excessive amount of processing

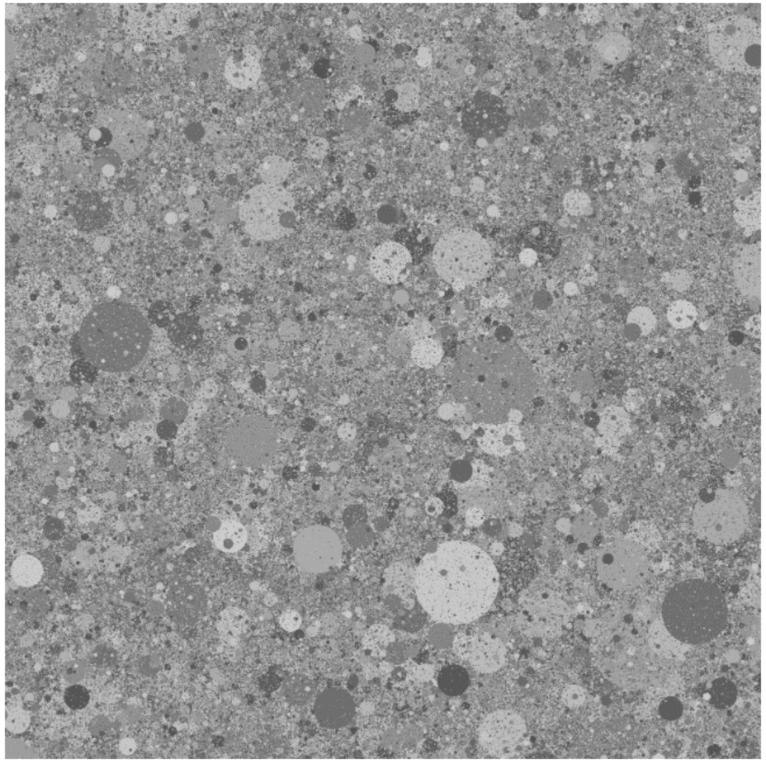


Photo V5 protocol | **Texture measurements** – DXOMARK's Dead Leaves Chart

What is on the chart:

- Markers at each corners, to automatically detect the target
- Dead Leaves pattern
- Gray patches around Dead Leaves of chosen reflectance to measure the camera OECF (also correct white balance) and **estimate the level of noise**
- Patterns in the corner with horizontal and vertical edges to measure an MTF
- 18% uniform gray background (mean reflectance for natural scenes)

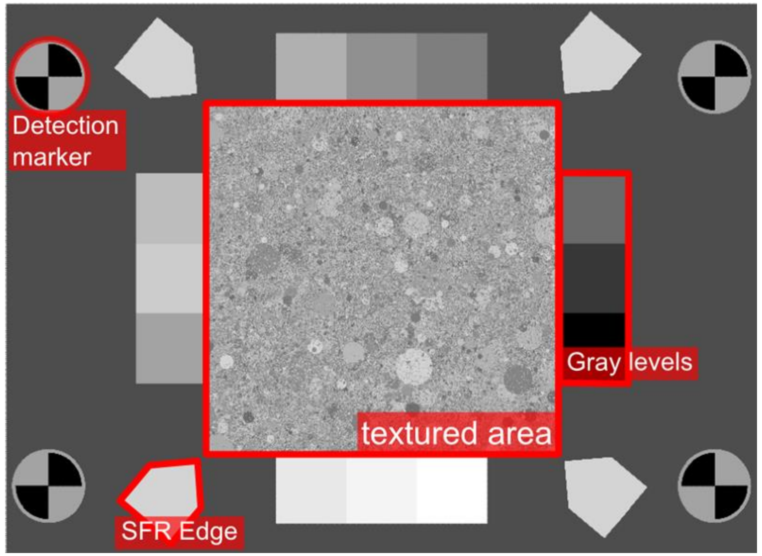


Photo V5 protocol | Texture measurements – DXOMARK's Dead Leaves Chart

To estimate the MTF, we model the imaging system as Linear Shift Invariant (LSI):

$$y(m,n) = (h * x)(m,n)$$

In Fourier, this gives:

$$Y(j,k) = H(j,k)X(j,k)$$

We also model an additive noise (measured on the gray patches):

$$Y(j,k) = H(j,k)X(j,k) + B(j,k)$$

The MTF of the system is the **radial mean of $\boldsymbol{H(j,k)}$**.

Our estimate compares the power spectrum densities of the studied image with the theoretical DeadLeaves pattern removing the influence of noise.

$$\boxed{|H(f)| = \sqrt{\frac{\phi_{YY}(f) - \phi_{BB}(f)}{\phi_{XX}(f)}}}$$

ϕ_{XX} is the power spectrum density of signal X, that's to say XX^*.

Photo V5 protocol | **Texture measurements** – DXOMARK's Dead Leaves Chart

MTF measurement workflow

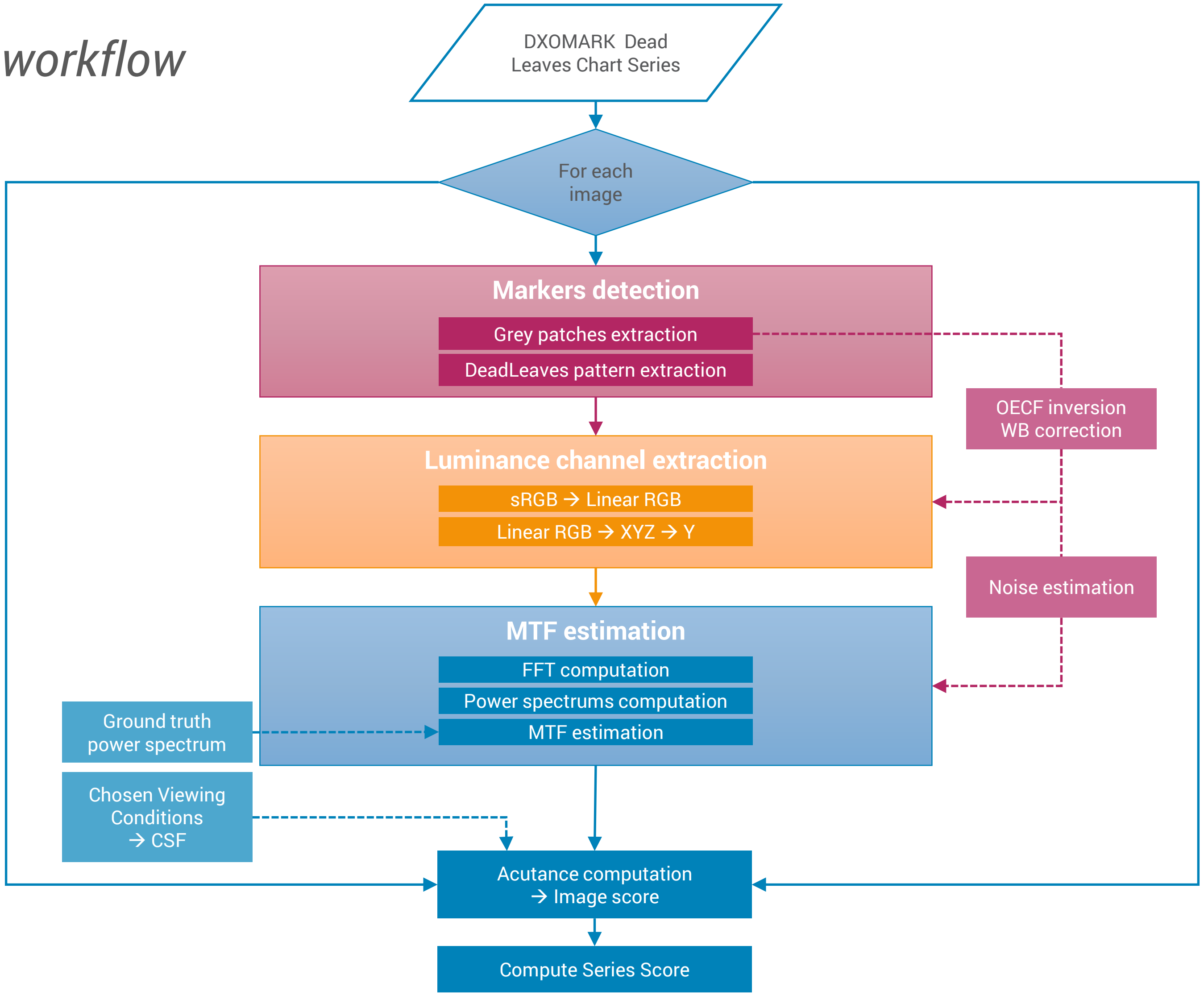
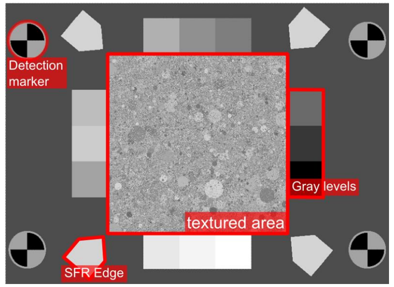


Photo protocol | **Texture & Noise** – Objective Measurements



Objective measurements of **texture acutance** are made on the **center** of the Dead Leaves chart, in both handheld and tripod.

2022 high-end smartphone camera

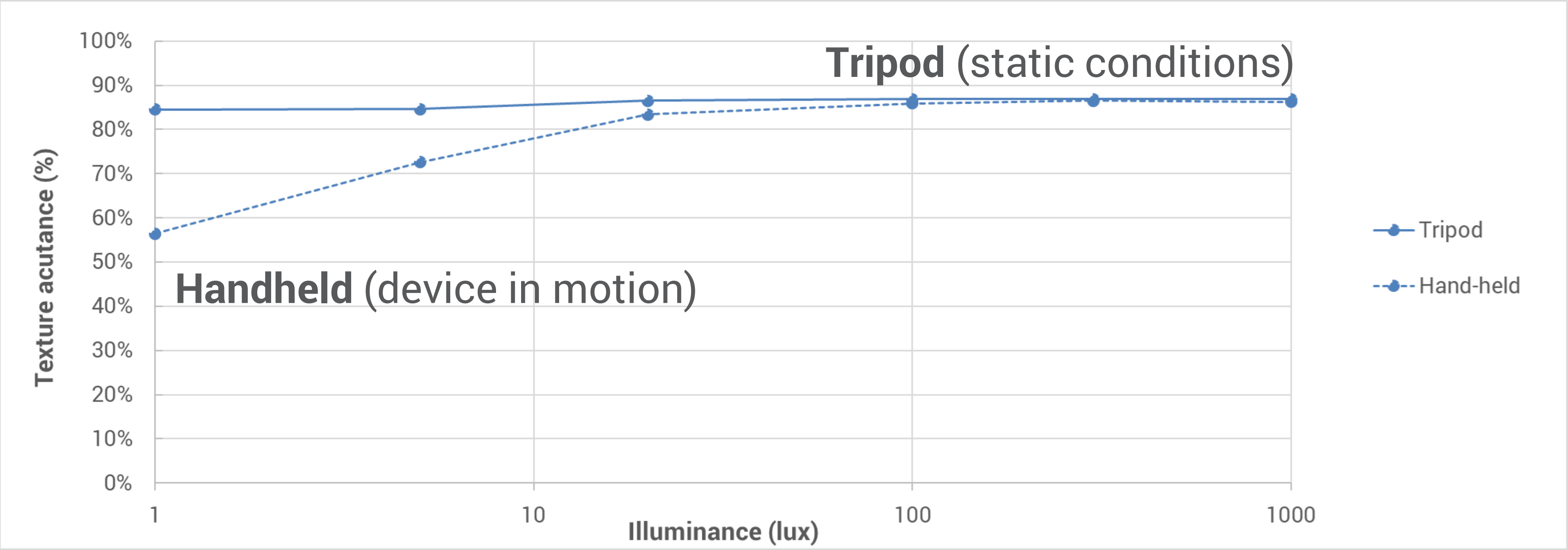
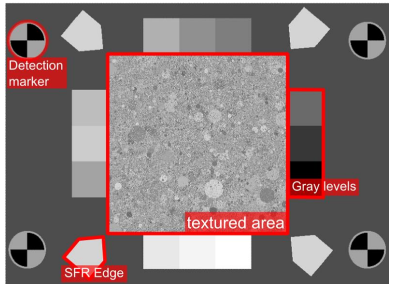


Photo protocol | **Texture & Noise** – Objective Measurements



Objective measurements of **texture acutance** are made on the **center** of the Dead Leaves chart, in both handheld and tripod.

2019 high-end smartphone camera (same manufacturer)

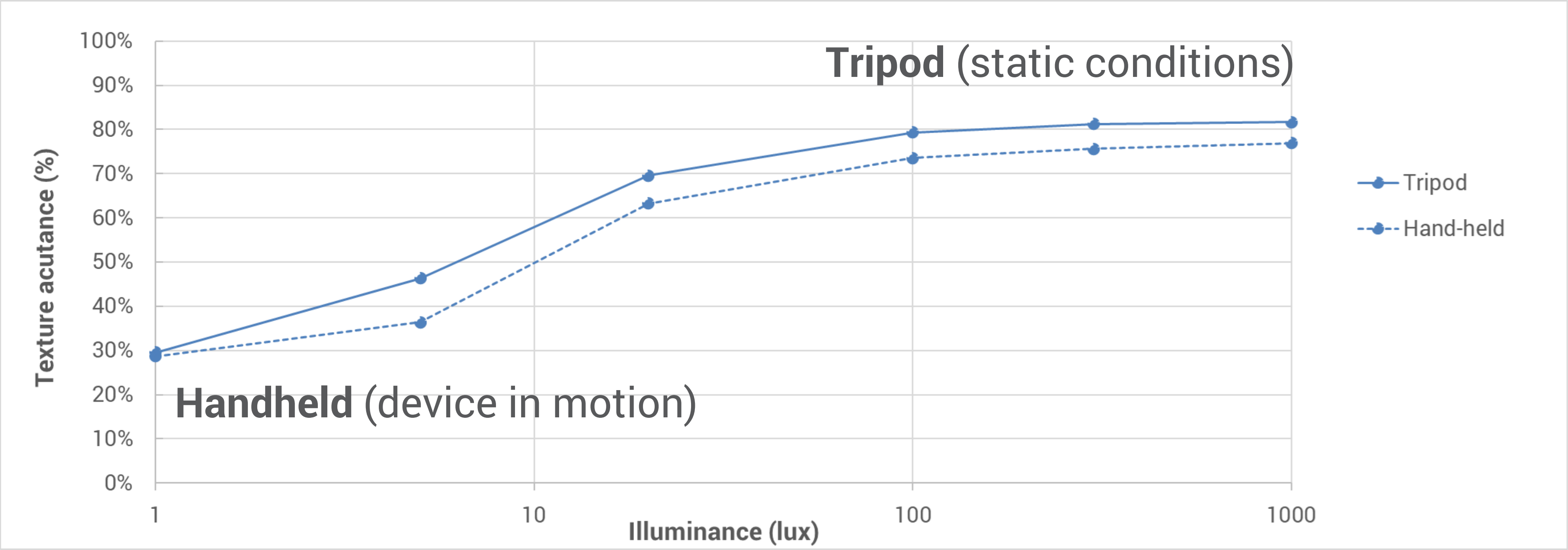


Photo V5 protocol | **Texture measurements** – limitations

This MTF computation relies purely on the **spectral content** of the image and an **estimation of noise on a uniform patch**:

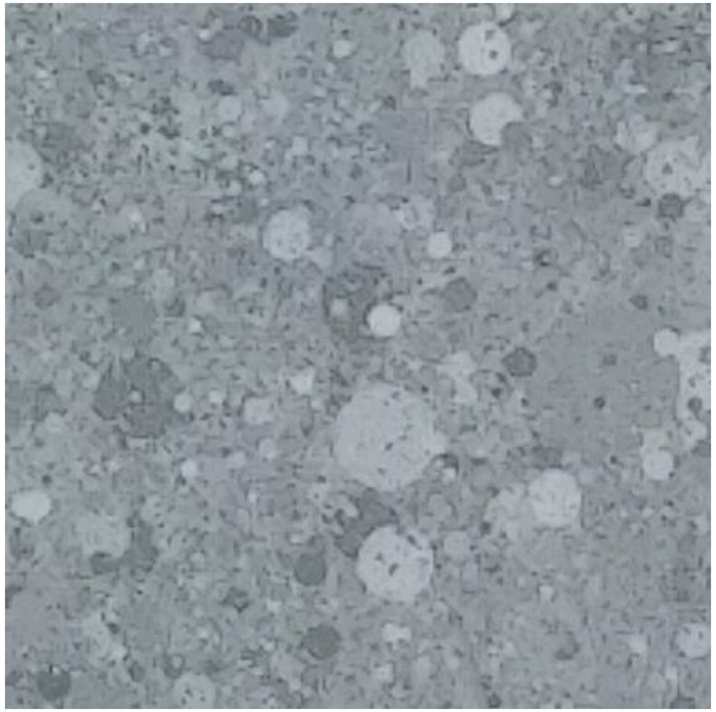
$$|H(f)| = \sqrt{\frac{\phi_{YY}(f) - \phi_{BB}(f)}{\phi_{XX}(f)}}$$

This estimation can be biased by today's "intelligent" denoising algorithms that can detect uniform areas where to apply strong denoising.

Hence, two images with same texture and level of noise on textured area but different noise levels on uniform grey patches might have different MTF, a behavior we would like to avoid.

Photo V5 protocol | **Texture measurements – limitations**

Other artifacts like blocking or sharpening might introduce “false” details and artificially improve texture scores



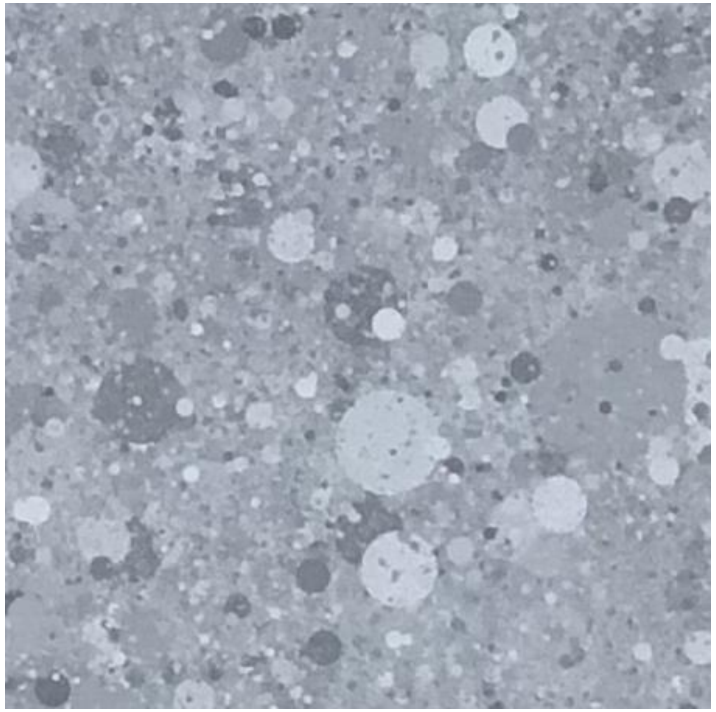
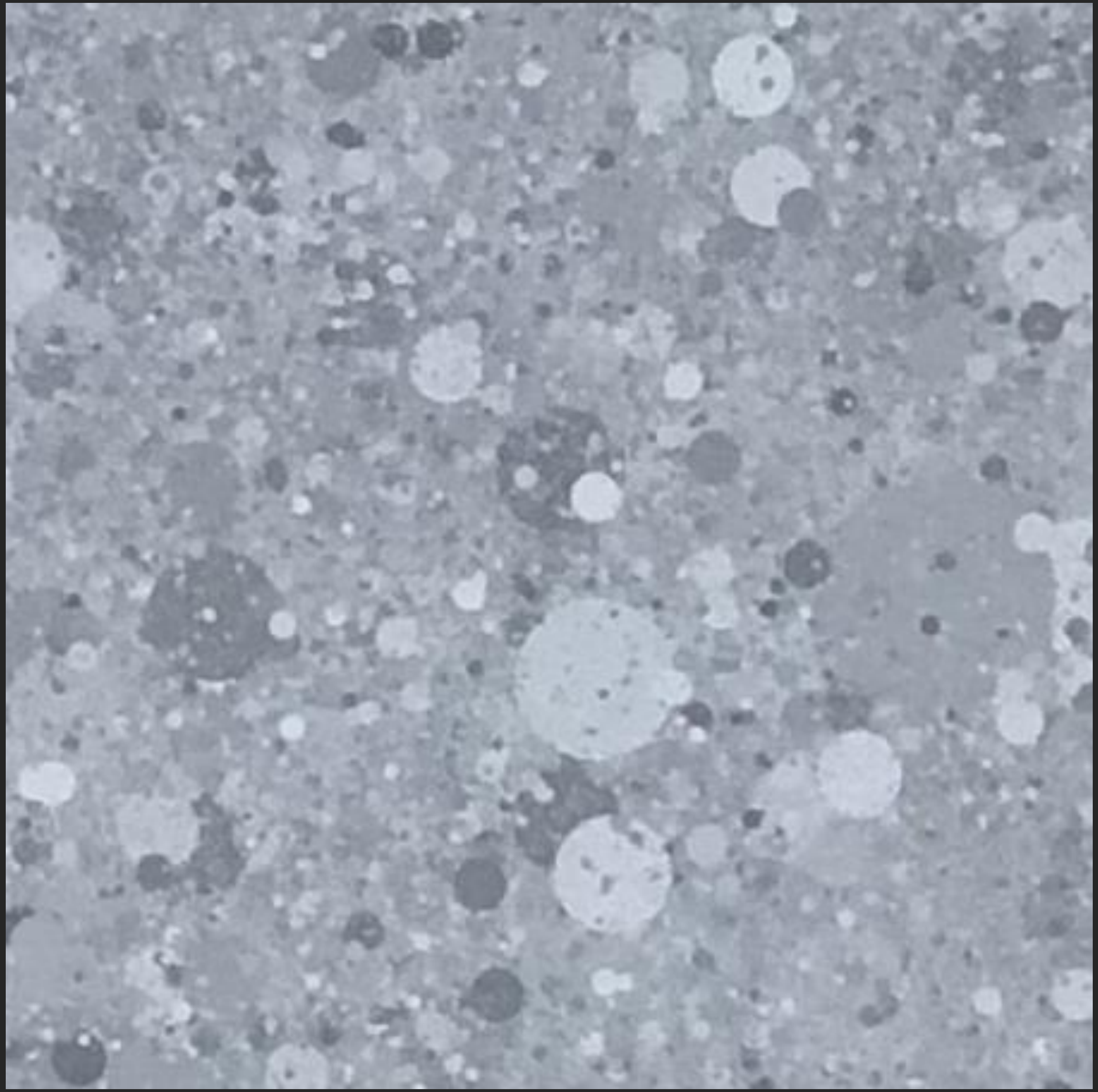


Image on the left introduces lots of artifacts while there are fewer details than picture on the right. These “false” details positively contribute to the MTF and bias the results.



Conclusions & a little history of camera metrology

What should you remember?

- Know how to read and request raw camera specifications for your applications:
  - Available range of settings: Exposure time(s), gain/ISO, aperture…
  - Methods for measuring: Noise/raw capture performance, sharpness and color
  - Departure from basic models e.g. nonlinearities / nonuniformities
  - Useful metrics for summarizing these measurement results

Sharpness	PSF, MTF Slanted edge method Checkerboard chart MTF10, MTF Ny/2
Noise	Transmissive gray patches Regimes of raw noise SNR curves, system gain SNR18%, *SNR10*
Color	Linear colorspaces, CIE-XYZ ColorChecker chart Linear color calibration Relationship with illuminant Metamerism, nonlinearities… Calibration quality

What should you remember?

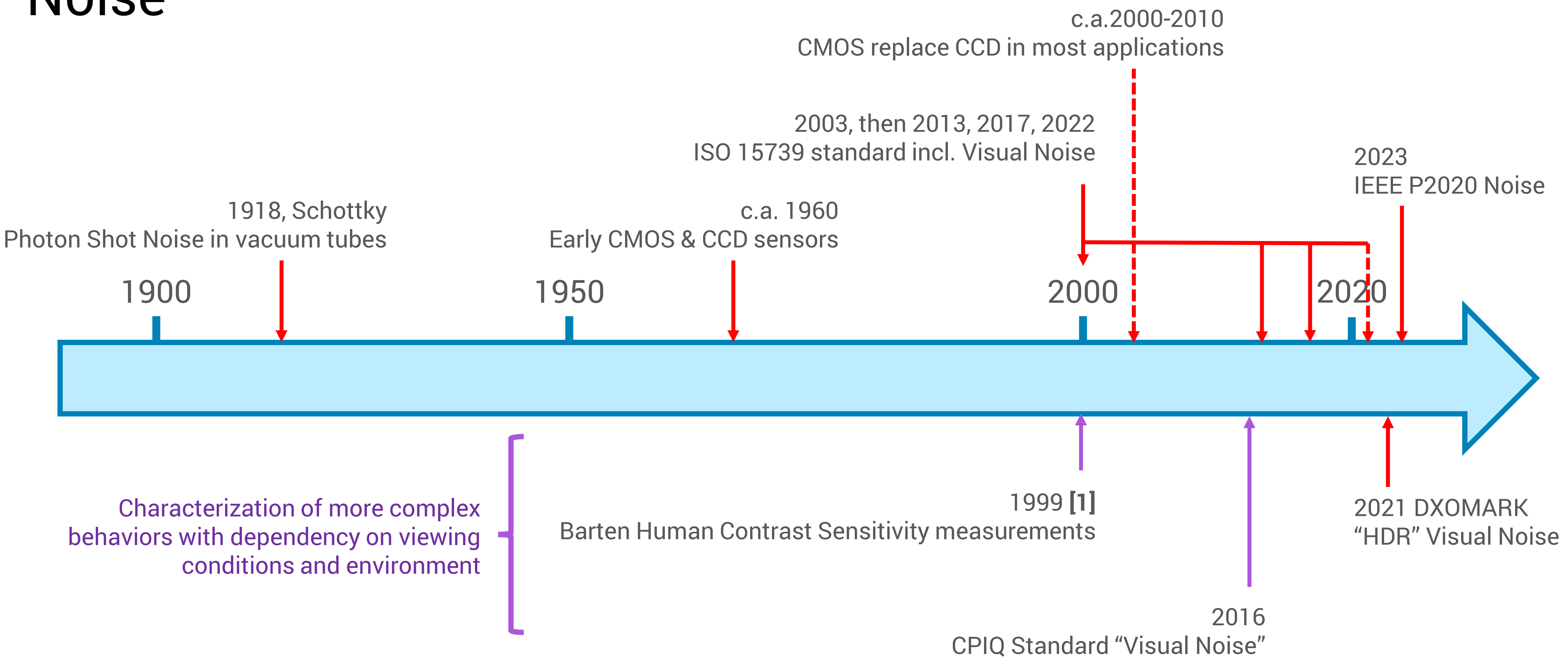
- Know how to do the same for tools & hardware that process RGB images:
  - Evaluation environment: viewing conditions, light levels, presence of motion,…
  - Metrics and methods for measuring: Perceptual noise, sharpness and color

Sharpness / Detail	Dead Leaves pattern Viewing conditions Edge Acutance Texture Acutance
Noise	Reflective gray patches Visual Noise Relationship with illuminance Tripod / Handheld
Color	ColorChecker chart CIELAB $\Delta a^* b^*$, ΔE User preferences Relationship with illuminant Color classes

What should you remember?

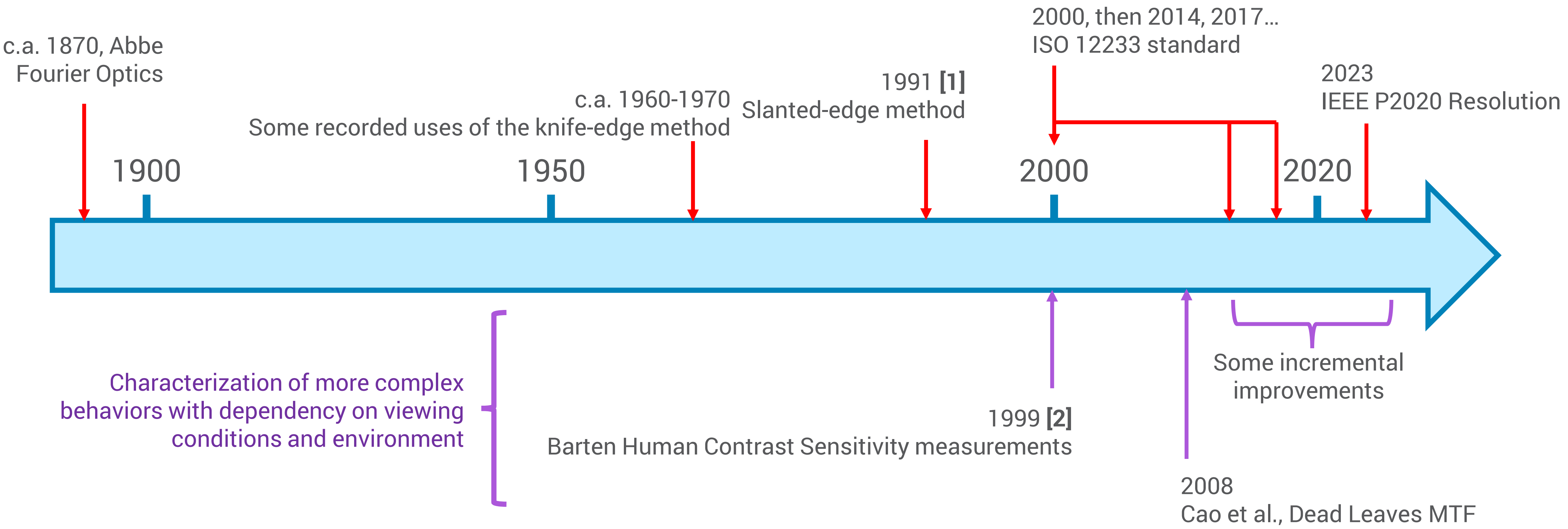
- More generally, have a rough idea of:
  - The landscape of camera performance evaluation methods
  - Their evolution, and why we sometimes need more complex methods

Noise



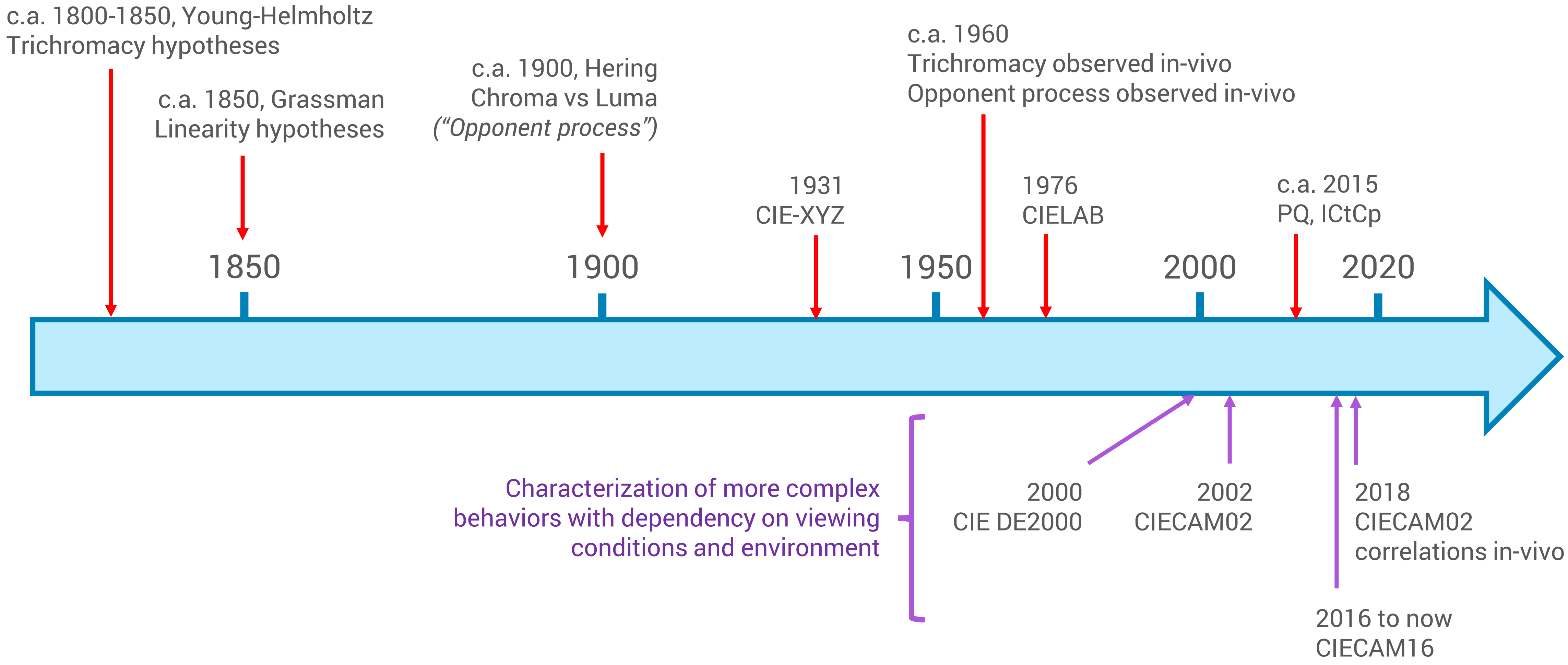
[1] Contrast Sensitivity of the Human Eye and Its Effects on Image Quality, Barten (1999)

Sharpness & Texture



[1] Characterizing digital image acquisition devices, Reichenbach et al. (1991)
[2] Contrast Sensitivity of the Human Eye and Its Effects on Image Quality, Barten (1999)

Color


c.a. 1800-1850, Young-Helmholtz
Trichromacy hypotheses
c.a. 1850, Grassman
Linearity hypotheses
c.a. 1900, Hering
Chroma vs Luma
("Opponent process")
1931
CIE-XYZ
c.a. 1960
Trichromacy observed in-vivo
Opponent process observed in-vivo
1976
CIELAB
c.a. 2015
PQ, ICtCp
1850
1900
1950
2000
2020
Characterization of more complex behaviors with dependency on viewing conditions and environment
2000
CIE DE2000
2002
CIECAM02
2018
CIECAM02
correlations in-vivo
2016 to now
CIECAM16

What you should take away from these

- Basics are sometimes >100 years old… and still in use today!
  - Basics of photometry, CIE-XYZ (1931),…
- More complex methods are not necessarily better
  - CIELAB (1976) still widely used despite its issues: it is simple and good enough in a lot of cases!
- State of the art vs. standardization (ISO, IEEE, CIE…)
  - Standards are slow to evolve – papers are published all the time!
  - A lot of details are present in the official standards but may not be replicated elsewhere:
    - Why a specific method is recommended / no longer recommended
    - Where the method came from, including mathematical proofs, required hypotheses, etc.
    - Which references you should read
  - Standards are maintained by their respective standardization groups
    - Add implementation notes explaining specific use cases or restrictions with new technologies
    - Correcting errors / typos
  - Standards are sometimes required in some industries (medical, automotive,…)

Interested in an internship or full-time job with us ?

https://dxomark-image-labs.breezy.hr

THANK YOU!

24-26, quai Alphonse le Gallo - F - 92100 Boulogne-Billancourt
www.dxomark.com